ACST3060 Assignment

Yunbae Chae 42809037

2023-10-17

Contents

Notes to markers & Usage of tutorial solution codes

I very much appreciate your time to mark my answers! I have skipped most obvious assumptions and included in the Appendix whatever I wasn't sure if it was okay to skip. Please note I have contained all answers within the 8 page limit (please count from page 4!), and all my codes are expressed in Appendix in the order of questions. *Regarding the codes in Questions 2.1, 2.2 and 2.4, I reviewed the Ed forum about whether it is plagiarism to use the tutorial answers, and I decided it would not be necessary to change the codes. I used the codes as what they are in the tutorial answers. My understanding is that using the codes in the solutions is not considered plagiarism. If this is wrong, could you please let me know and I will provide my own version of the codes!

I hope you enjoy marking, and please note I have prepared my answers for a potential full mark.

Once again, thank you for your time!

Kind regards, Yunbae Chae

Question 1.1

$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S|N]] = \mathbb{E}[n \cdot \mathbb{E}[X_1]] = \mathbb{E}[N] \cdot \mathbb{E}[X] = \frac{r(1-p)}{p} \cdot \frac{\gamma}{\lambda}$

$\mathbb{V}[S] = \mathbb{E}[X]^2 \cdot \mathbb{V}[N] + \mathbb{E}[N] \cdot \mathbb{V}[X] = \frac{\gamma^2}{\lambda^2} \cdot \frac{r(1-p)}{p^2} + \frac{r(1-p)}{p} \cdot \frac{\gamma}{\lambda^2} = \frac{\gamma r(1-p)[\gamma+p]}{p^2\lambda^2}$

$$CV = \frac{\sqrt{\mathbb{V}[S]}}{\mathbb{E}[S]} = \frac{\sqrt{\gamma r(1-p)[\gamma+p]}}{p\lambda} \cdot \frac{p\lambda}{\gamma r(1-p)} = \sqrt{\frac{\gamma+p}{\gamma r(1-p)}}$$

Parameter features to consider:

γ: A higher shape parameter makes the distribution of individual values more peaked and less variable.

p: As p approaches 1, each Bernoulli trial is almost certain to result in success and therefore the number of failure N is close to zero.

r: A larger r generates a larger sample size and by the CLT, the variability will almost certainly reduce.

$$\lim_{\gamma\to\infty} \sqrt{\frac{\gamma+p}{\gamma r(1-p)}} = \lim_{\gamma\to\infty} \sqrt{\frac{1}{r(1-p)} + \frac{1}{\gamma r} \cdot \frac{1}{\frac{1}{p}-1}} = \sqrt{\frac{1}{r(1-p)}}$$

As γ gets bigger, the CV reduces towards the lower bound $\sqrt{\frac{1}{r(1-p)}}$. This is because a higher shape parameter makes the distribution more peaked and sharp around the mean.

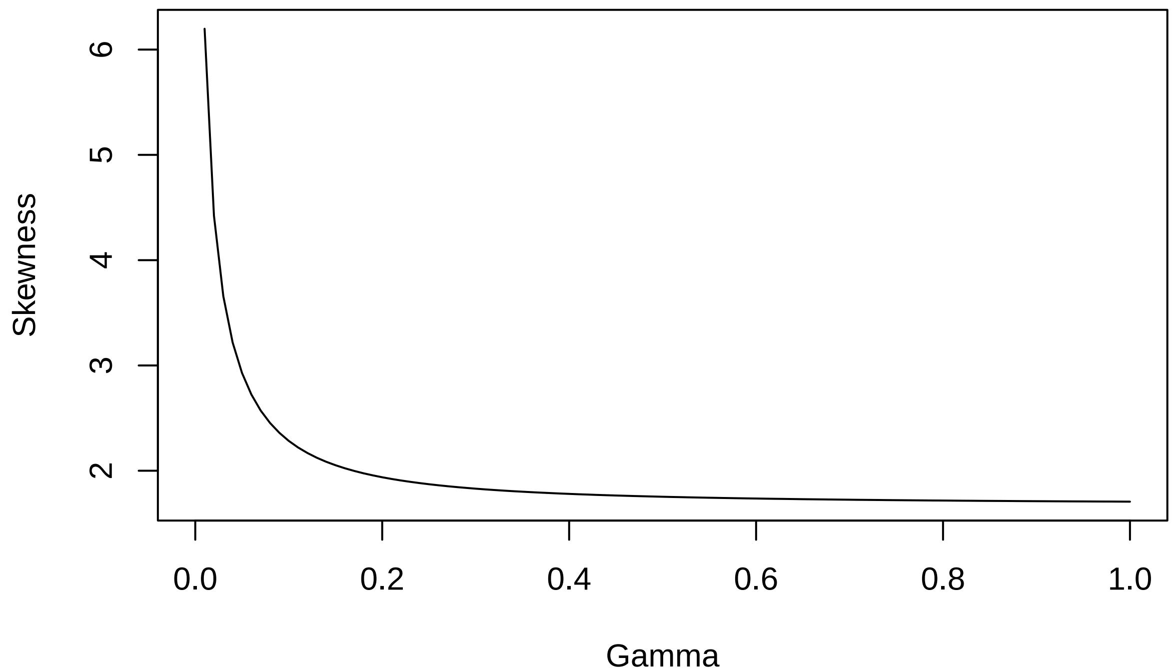
$$\lim_{p\to1} \sqrt{\frac{1}{r(1-p)} + \frac{1}{\gamma r} \cdot \frac{1}{\frac{1}{p}-1}} \to \infty$$

As p approaches 1, the CV goes to infinity. This is because it is almost certain to have no fail before the r number of success. As expectation of N approaches zero while variability still exists for S, the CV must approach infinity. $\lim_{r\to\infty} \sqrt{\frac{\gamma+p}{\gamma r(1-p)}} = 0$

As r approaches infinity, CV approaches zero. This is because a very large number of samples would result in the sample mean to become ever closer to the population mean.

Question 1.2

My student ID: 42809037: $r = 1.43,\ p = 0.12,\ \lambda = 1.77$



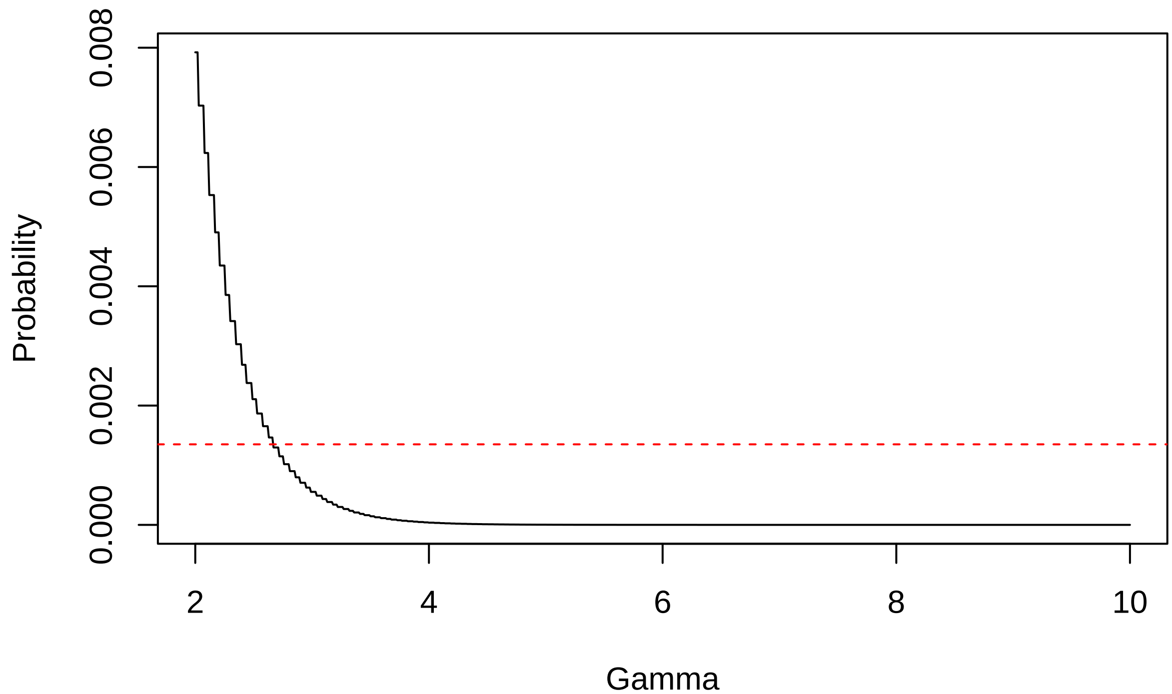
For all values $\gamma > 0$, the Skewness is at least 1.6759015. Intuitively, with all other parameters held constant, a higher γ leads to the heavier probabilities in bigger claim sizes. In other words, the S with a higher γ would now consist of bigger claim sizes. Higher claims that were once on the right-skew side are now on the middle. S has become more symmetrical; The skewness of S has decreased, with a higher γ value.

Question 1.3

The probabilities of $\gamma = 0.1,\ 2,\ 100$ are 0.8262201, 0.0079231, 0.

The normal approximation of: $Pr\left[\frac{S-\mathbb{E}[S]}{\sqrt{\mathbb{V}[S]}} > 3\right]$=0.0013499

This lies between $\gamma = 2$ and $\gamma = 100 \in (0.0079231,\ 0)$



What the results show is that the γ value roughly around 2.7 coincides with the probability estimated using the Normal assumption for S. This is the point where the γ value that determines the shape of the compound distribution S is closest to the bell shape which is the main characteristic of a Normal Distribution.

In addition, different gamma values change the value of $\mathbb{E}[S] + 3\sqrt{\mathbb{V}[S]}$ significantly. Since S is right-skewed, the area to the right side of each vertical lines ($x = \mathbb{E}[S] + 3\sqrt{\mathbb{V}[S]}$) under the PDF curve must decrease as γ increases, as observed: (0.8262201, 0.0079231, 0, for $\gamma = 0.1, 2, 100$).

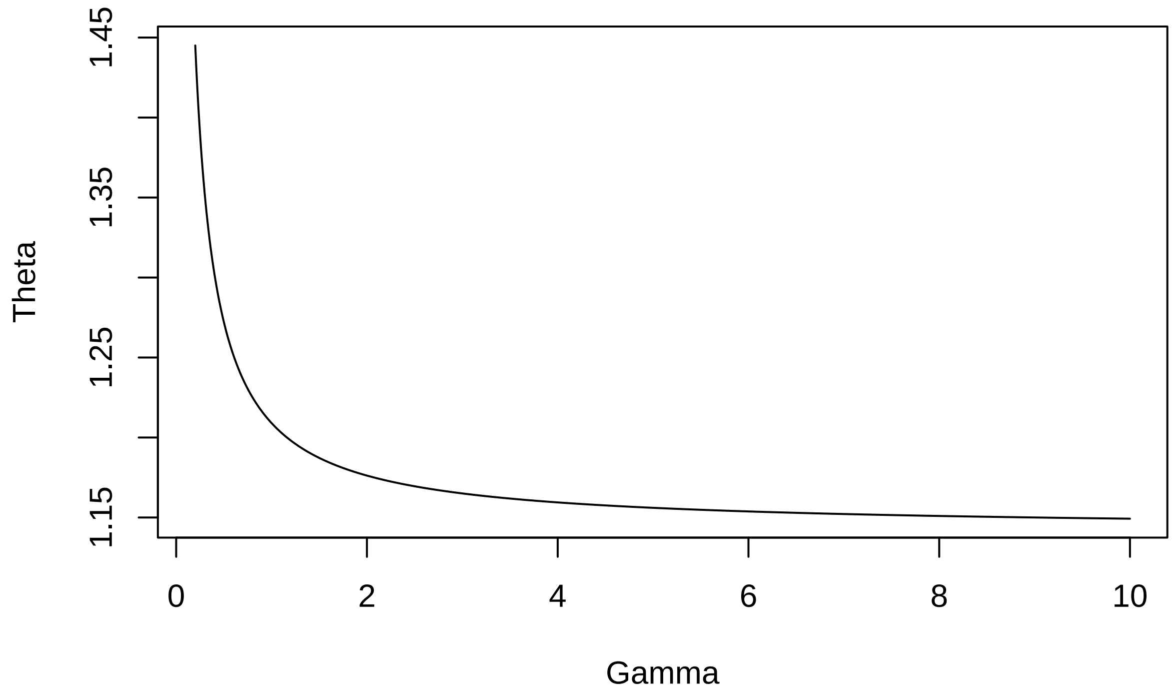
Question 1.4

$$Pr[\Pi(S) > S] \geq 0.9 \Rightarrow Pr[(1+\theta)\mathbb{E}[S] > S] \geq 0.9$$

$$Pr\Big[\frac{S - \mathbb{E}[S]}{\sqrt{\mathbb{V}[S]}} < \frac{\theta\mathbb{E}[S]}{\sqrt{\mathbb{V}[S]}}\Big] \geq 0.9 \Rightarrow \frac{\theta\mathbb{E}[S]}{\sqrt{\mathbb{V}[S]}} \geq \Phi^{-1}(0.9)$$

$\Rightarrow \theta \geq \Phi^{-1}(0.9) \cdot \frac{\sqrt{\mathbb{V}[S]}}{\mathbb{E}[S]}$=1.2721478 (Minimum loading 127.2147806 %)

Question 1.5



θ is a decreasing function of γ. This must be due to the increasing γ concentrates probability of claims around the mean, reducing the uncertainty in the range of the magnitude of claims. The decreasing uncertainty would result in requiring less loading to secure 90% probability of making a profit. The function converges to the lower bound that is >0, which is the absolute minimum loading to ensure the 90% probability of profit.

Question 1.6

Using given Skew (more reasoning in Appendix),

$Skew[S] = \frac{2p^2+3\gamma p+\gamma^2(2-p)}{\sqrt{r\gamma(1-p)}(p+\gamma)^{3/2}}$=1.7529113

$\alpha = \frac{4}{Skew[S]^2} = \frac{4}{1.7529113^2} = 1.3017875$

$\beta = \sqrt{\frac{\alpha}{\mathbb{V}[S]}}$ ==0.3880024

$k = \mathbb{E}[S] - \frac{\alpha}{\beta}$=-0.3927669

$Pr[S < (1+\theta)\mathbb{E}[S]] \geq 0.9 \Rightarrow Pr[Y + k < (1+\theta) \cdot (\frac{\alpha}{\beta} + k)] \geq 0.9$

$$\Rightarrow Pr[Y < (1+\theta) \cdot (\frac{\alpha}{\beta} + k) - k] \geq 0.9 \Rightarrow (1+\theta) \cdot (\frac{\alpha}{\beta} + k) - k \geq F_Y^{-1}(0.9) = 7.2385$$

$\Rightarrow \theta \geq \frac{7.2385+k}{\frac{\alpha}{\beta}+k} - 1$=1.3109246

The relative error from the original θ is: $|\frac{1.3109246-1.2721478}{1.2721478}| = 0.0304814$

The relative error of the 'translated Gamma' approximation for the profit loading θ is 3.05%. This is a fairly small error. Limited to the specific context of $\gamma = 0.5$, and other given parameters, this approximation seems to be a good one, because of the small error rate 3.05%. However, as the question stated, it can never be guaranteed the approximation will always be good for other values of Gamma. Therefore, it should be emphasised that the 'translated Gamma' approximation in 'this specific parameter set up', is good.

Question 1.7

The insurer would only find it favorable to enter the re-insurance arrangement if the expected utility of paying the premium and receiving the recovery is greater or at least equal to the utility of not doing anything:
$\Rightarrow u(w_0) \leq \mathbb{E}[u(w_0 - 2\mathbb{E}[Z] + Z)] \Rightarrow -e^{-Bw_0} \leq \mathbb{E}[-e^{-B(w_0 - 2\mathbb{E}[Z]+Z)}]$

$\Rightarrow -e^{-Bw_0} \leq -e^{-Bw_0} \cdot e^{2B\mathbb{E}[Z]} \cdot \mathbb{E}[e^{-BZ}] \Rightarrow 1 \geq e^{2B\mathbb{E}[Z]} \cdot \mathbb{E}[e^{-B(S-Y)}]$

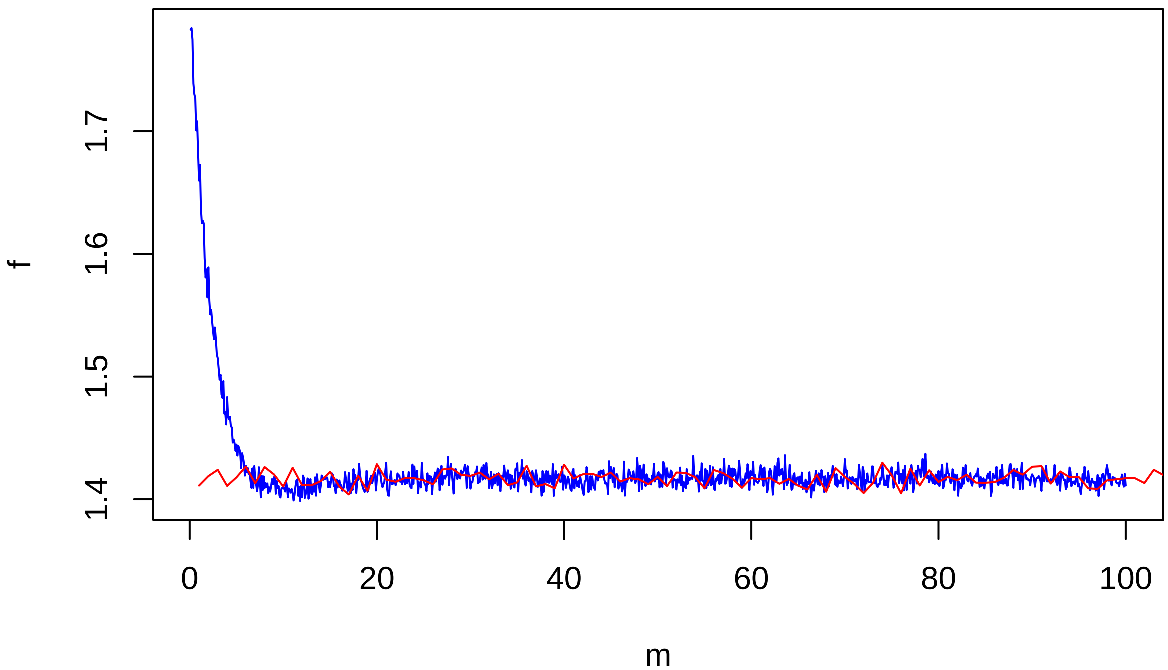
Since S and Y are identical for $S \leq m$ and independent for $S > m$, $\mathbb{E}[e^{-B(S-Y)}] = \mathbb{E}[e^{-BS}] \cdot \mathbb{E}[e^{BY}]$.

$1 \geq e^{2B\mathbb{E}[Z]} \cdot \mathbb{E}[e^{-BS}] \cdot \mathbb{E}[e^{BY}] \Rightarrow \mathbb{E}[e^{BS}] \geq e^{2B\mathbb{E}[Z]} \cdot \mathbb{E}[e^{-BS}] \cdot \mathbb{E}[e^{BS}] \cdot \mathbb{E}[e^{BY}]$

$$\therefore \mathbb{E}[e^{BY}] \cdot e^{2B\mathbb{E}[Z]} \leq \mathbb{E}[e^{BS}]$$

The above is the condition that the insurer finds it favourable to enter the reinsurance arrangement.

Question 1.8



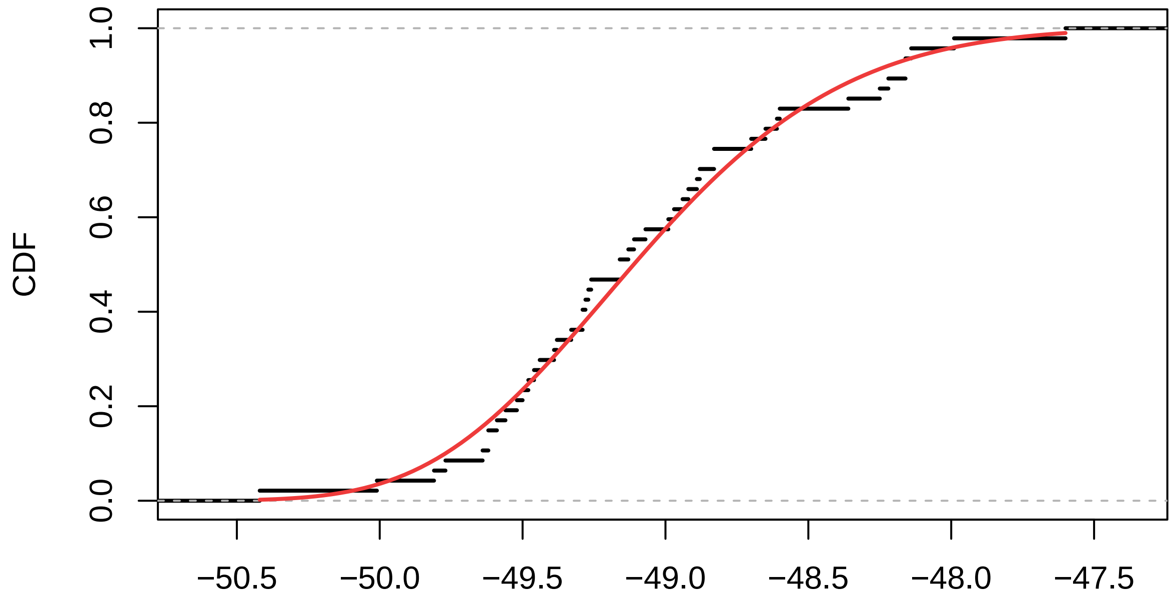

```
##  [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9
## [20] 2.0 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8
## [39] 3.9 4.0 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8 4.9 5.0 5.1 5.2 5.3 5.4 5.5 5.6 5.7
## [58] 5.8 5.9 6.0 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 7.2
```

The above array of m values are all entries of m whenever $\mathbb{E}[e^{BY}] \cdot e^{2B\mathbb{E}[Z]}$ (blue) is greater than $\mathbb{E}[e^{BS}]$ (red), and this is the range of the retention limits that are unfavorable for the insurer. In this simulation, 7.3 is the first retention limit where the $\mathbb{E}[e^{BY}] \cdot e^{2B\mathbb{E}[Z]} \leq \mathbb{E}[e^{BS}]$ holds. Running the simulation 100 times, the average value of the retention limit m is 7.583. From m = 7.583 in average, for $m \in (0, 100)$, 93% of the time, the insurer is subject to the favorable condition. This rate is 99.3% for $m \in (0, 1000)$. It is a strong evidence that the range of m with which the insurer would be willing to purchase the reinsurance is $m \geq 7.583$, because $\mathbb{E}[e^{BY}] \cdot e^{2B\mathbb{E}[Z]} \leq \mathbb{E}[e^{BS}]$ holds for $m \geq 7.583$, 93.1% and 99.3% of the time for $m \in (0, 100)$ and $m \in (0, 1000)$.

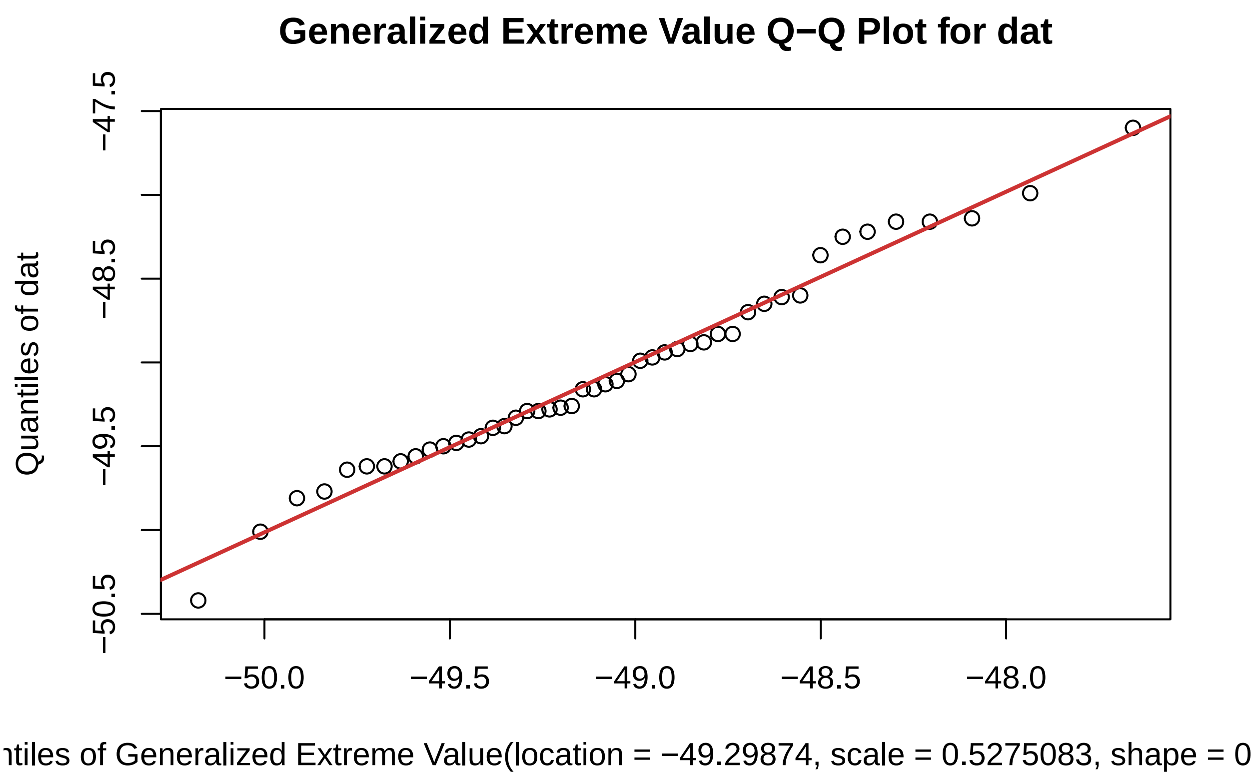
Question 2.1

The estimates $\hat{\mu}$, $\hat{\sigma}$, $\hat{\xi}$ respectively, are -49.292556, 0.5320409, -0.1853549 (Codes in Appendix). The negative ξ means the distribution is of the reversed Weibull type. It has a heavy left tail and a lighter right tail. Under this set up, the probability of observing extremely fast running times is much smaller than that of observing slower running times.

Question 2.2



CDF
0.0
0.2
0.4
0.6
0.8
1.0
−50.5
−50.0
−49.5
−49.0
−48.5
−48.0
−47.5



Generalized Extreme Value Q−Q Plot for dat
Quantiles of dat
−50.5
−49.5
−48.5
−47.5
−50.0
−49.5
−49.0
−48.5
−48.0
ntiles of Generalized Extreme Value(location = −49.29874, scale = 0.5275083, shape = 0

The theoretical CDF shows a very good match against the empirical CDF. The theoretical CDF therefore seems to be a good approximation for the distribution of the observed data. The QQ plot intertwining with the straight line indicates the distribution of the data is well approximated by the theoretical distribution.

Question 2.3

Since the distribution is of the reversed Weibull type, it is subject to the upper bound $\mu - \frac{\sigma}{\xi}$.

Using estimates, $\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}$ =-46.0941939

The absolute best time that can ever be achieved is 46.0941939 seconds.

The probability that the world record of 47.60s will be beaten within the next 100 years = 1 - NOT beating the record at all for 100 years:

$$1 - P(X \leq -47.60)^{100} = 1 - G_{\hat{\mu},\ \hat{\sigma},\ \hat{\xi}}(-47.60)^{100} = 1 - exp\Big[- (1 + \hat{\xi} \cdot \tfrac{-47.60-\hat{\mu}}{\hat{\sigma}})^{-\frac{1}{\hat{\xi}}}\Big]^{100}$$

$$= 1 - exp\Big[- (1 - 0.1853549 \cdot \tfrac{-47.60+49.2925560}{0.5320409})^{-\frac{1}{-0.1853549}}\Big]^{100} = 0.5587969$$

Question 2.4

The estimates $\hat{\mu},\ \hat{\sigma},\ \hat{\xi}$ in this set up respectively, are -49.3884138, 0.451613, -0.2377146.

the upper bound $\mu - \frac{\sigma}{\xi}$ this time is, using the same $\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}$, 46.0941939

Using same method, $= 1 - exp\Big[- (1 - 0.2377146 \cdot \tfrac{-47.60+49.3884138}{0.451613,})^{-\frac{1}{-0.2377146}}\Big]^{100}$

=6.5747319×10^{-4}

Question 2.5

The two counterparts are extremely different. The scores from 1990 and onwards suggest it is basically impossible to beat the world record for the next 100 years (P=6.5747319×10^{-4}). Since the 1990 and onwards record is based on the presense of the random doping test, it seems very obvious that the players stopped using drugs and their running times were significantly affected. It seems highly evident, based on the statistic findings, that the players were using drugs prior to 1990. The 1985 world record was most likely obtained in an illegal manner.

Appendix

Background reasonings

Question 1.1

Moments of the Generalised Negative Binomial Distribution Below is the iteration of the process of finding $\mathbb{E}$ and $\mathbb{V}$ allowing for r to be a non-integer. They are exactly same with what is on the formula sheet.. Please ignore if irrelavant:

$$\mathbb{E}[N] = \sum_{k=0} k \cdot Pr[N = k] = \sum_{k=1} k \cdot Pr[N = k]$$

$$= \sum_{k=1} \frac{\Gamma(k + r)}{(k - 1)! \cdot \Gamma(r)}(1 - p)^k p^r = \sum_{k=1} \frac{\Gamma(k + r)}{\Gamma(k) \cdot \Gamma(r)}(1 - p)^k p^r$$

$$= \frac{\Gamma(1 + r)}{\Gamma(1) \cdot \Gamma(r)}(1 - p)^1 p^r + \frac{\Gamma(2 + r)}{\Gamma(2) \cdot \Gamma(r)}(1 - p)^2 p^r + \frac{\Gamma(3 + r)}{\Gamma(3) \cdot \Gamma(r)}(1 - p)^3 p^r + \cdots$$

$$= \frac{(1 - p) \cdot r}{p}\Big[\frac{\Gamma(1 + r)}{\Gamma(1) \cdot \Gamma(r + 1)}(1-p)^0 p^{r+1} + \frac{\Gamma(2 + r)}{\Gamma(2) \cdot \Gamma(r + 1)}(1-p)^1 p^{r+1} + \frac{\Gamma(3 + r)}{\Gamma(3) \cdot \Gamma(r + 1)}(1-p)^2 p^{r+1} + \cdots\Big]$$

The sum of all probabilities of a P.M.F is equal to 1:

$$\sum_{x=0}^{\infty} \frac{\Gamma(x + r)}{x!\Gamma(r)}(1 - p)^x p^r = \frac{\Gamma(r)}{\Gamma(1)\Gamma(r)}(1 - p)^0 p^r + \frac{\Gamma(1 + r)}{\Gamma(2)\Gamma(r)}(1 - p)^1 p^r + \frac{\Gamma(2 + r)}{\Gamma(3)\Gamma(r)}(1 - p)^2 p^r + \cdots = 1$$

Manipulate this by substituting r=z+1:

$$= \frac{\Gamma(z + 1)}{\Gamma(1)\Gamma(z + 1)}(1 - p)^0 p^{z+1} + \frac{\Gamma(z + 2)}{\Gamma(2)\Gamma(z + 1)}(1 - p)^1 p^{z+1} + \frac{\Gamma(z + 3)}{\Gamma(3)\Gamma(z + 1)}(1 - p)^2 p^{z+1} + \cdots = 1$$

This fits in the previous $\mathbb{E}[N]$ setting:

$$\mathbb{E}[N] = \frac{(1 - p) \cdot r}{p}\Big[\frac{\Gamma(z + 1)}{\Gamma(1)\Gamma(z + 1)}(1-p)^0 p^{z+1} + \frac{\Gamma(z + 2)}{\Gamma(2)\Gamma(z + 1)}(1-p)^1 p^{z+1} + \frac{\Gamma(z + 3)}{\Gamma(3)\Gamma(z + 1)}(1-p)^2 p^{z+1} + \cdots\Big]$$

$$\therefore \mathbb{E}[N] = \frac{(1-p)\cdot r}{p}$$

$$\mathbb{V}[N] = \mathbb{E}[N(N-1)] + \mathbb{E}[N] - \mathbb{E}[N]^2$$

$$\mathbb{E}[N(N-1)] = \sum_{k=2} \frac{\Gamma(k+r)}{\Gamma(k-1)\cdot\Gamma(r)}(1-p)^k p^r$$

$$= \frac{\Gamma(2+r)}{\Gamma(1)\cdot\Gamma(r)}(1-p)^2 p^r + \frac{\Gamma(3+r)}{\Gamma(2)\cdot\Gamma(r)}(1-p)^3 p^r + \frac{\Gamma(4+r)}{\Gamma(3)\cdot\Gamma(r)}(1-p)^4 p^r + \cdots$$

$$= \frac{(1-p)^2\cdot r\cdot(r+1)}{p^2}\Big[\frac{\Gamma(2+r)}{\Gamma(1)\cdot\Gamma(r+2)}(1-p)^0 p^{r+2} + \frac{\Gamma(3+r)}{\Gamma(2)\cdot\Gamma(r+2)}(1-p)^1 p^{r+2} + \frac{\Gamma(4+r)}{\Gamma(3)\cdot\Gamma(r+2)}(1-p)^2 p^{r+2} + \cdots\Big]$$

The sum of all probabilities of a P.M.F is equal to 1:

$$\sum_{x=0}^{\infty} \frac{\Gamma(x+r)}{x!\Gamma(r)}(1-p)^x p^r = \frac{\Gamma(r)}{\Gamma(1)\Gamma(r)}(1-p)^0 p^r + \frac{\Gamma(1+r)}{\Gamma(2)\Gamma(r)}(1-p)^1 p^r + \frac{\Gamma(2+r)}{\Gamma(3)\Gamma(r)}(1-p)^2 p^r + \cdots = 1$$

Manipulate this by substituting r=z+2:

$$= \frac{\Gamma(z+2)}{\Gamma(1)\Gamma(z+2)}(1-p)^0 p^{z+2} + \frac{\Gamma(z+3)}{\Gamma(2)\Gamma(z+2)}(1-p)^1 p^{z+2} + \frac{\Gamma(z+4)}{\Gamma(3)\Gamma(z+2)}(1-p)^2 p^{z+2} + \cdots = 1$$

By fitting 1 into the $[\cdots]$,

$$\mathbb{E}[N(N-1)] = \frac{(1-p)^2\cdot r\cdot(r+1)}{p^2}$$

$$\mathbb{V}[N] = \frac{(1-p)^2\cdot r\cdot(r+1)}{p^2} + \frac{(1-p)\cdot r}{p} - \frac{(1-p)^2\cdot r^2}{p^2}$$

$$= \frac{(1-p)[(1-p)(r^2+r)+rp-(1-p)r^2]}{p^2}$$

$$= \frac{r(1-p)}{p^2}$$

Moments of X

$$\mathbb{E}[X] = \frac{\gamma}{\lambda}$$

$$\mathbb{V}[X] = \frac{\gamma}{\lambda^2}$$

Question 1.6

Translated Gamma

$$\mathbb{E}[S] = \frac{\alpha}{\beta} + k$$

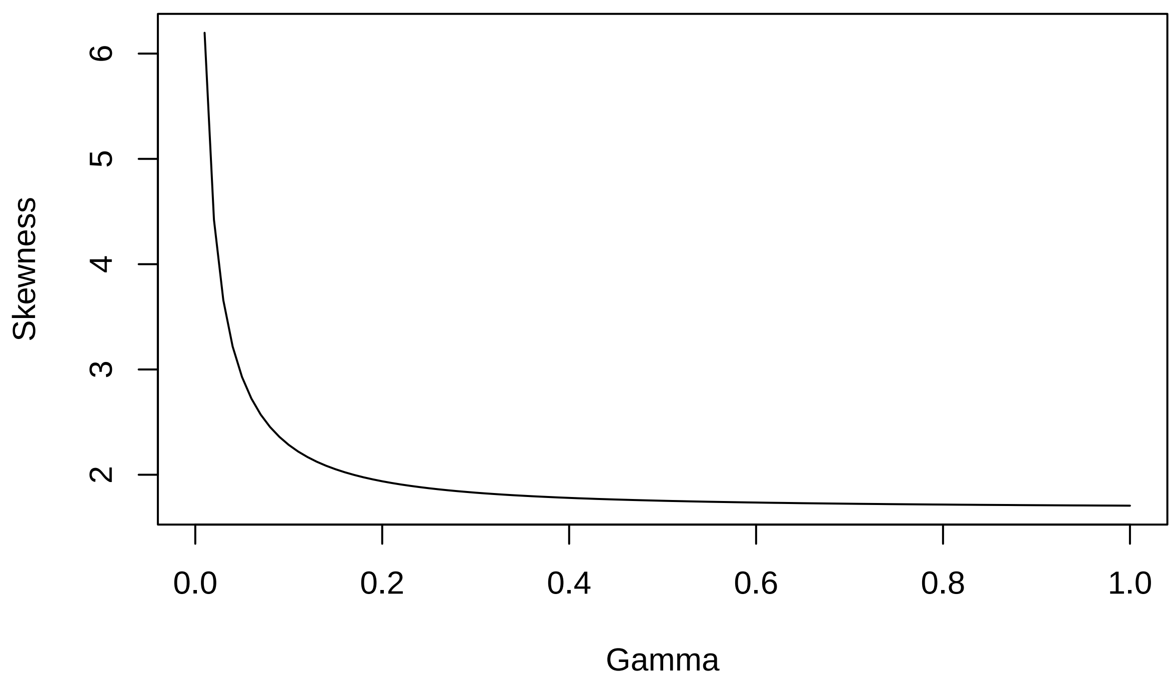
$$\mathbb{V}[S] = \frac{\alpha}{\beta^2}$$

$$Skew[S] = \frac{2}{\sqrt{\alpha}}$$

$$\alpha = \frac{4}{Skew[S]^2}$$

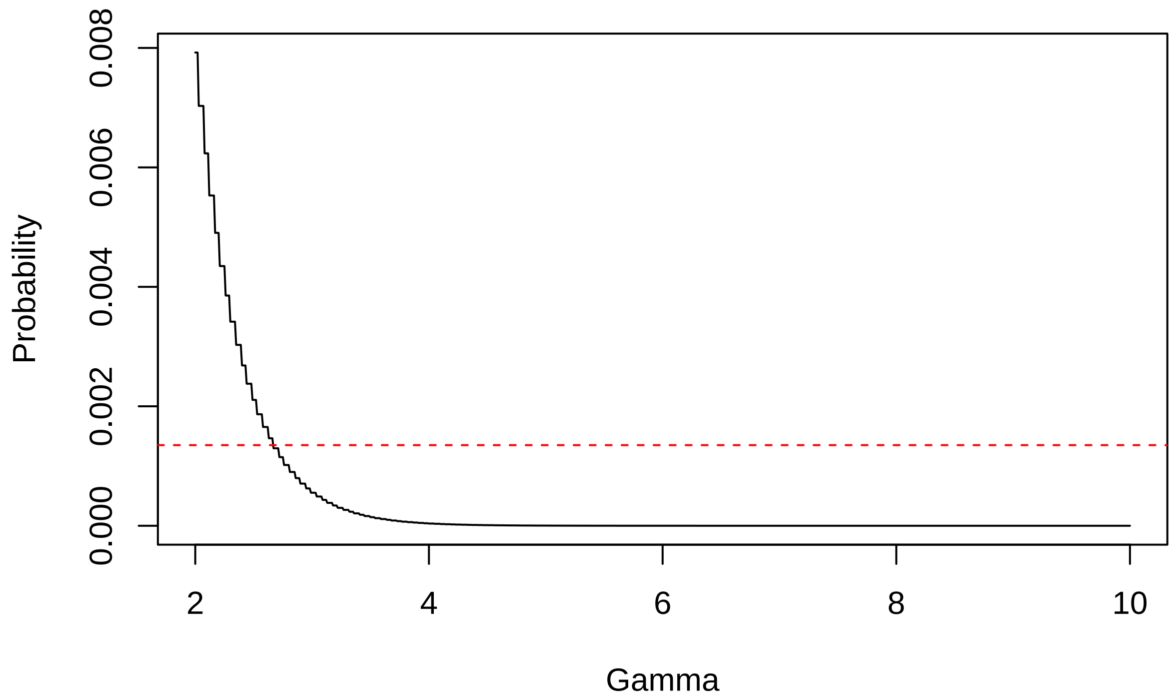
R Codes

```
#Part One - Q2
r <- 1.43
p <- 0.12
lambda <- 1.77
gamma <- seq(0,1,by=0.01)

Skew <- (2*p^2+3*gamma*p+gamma^2*(2-p))/(sqrt(r*gamma*(1-p))*(p+gamma)^1.5)
plot(gamma, Skew, type="l", xlab="Gamma", ylab="Skewness")
```

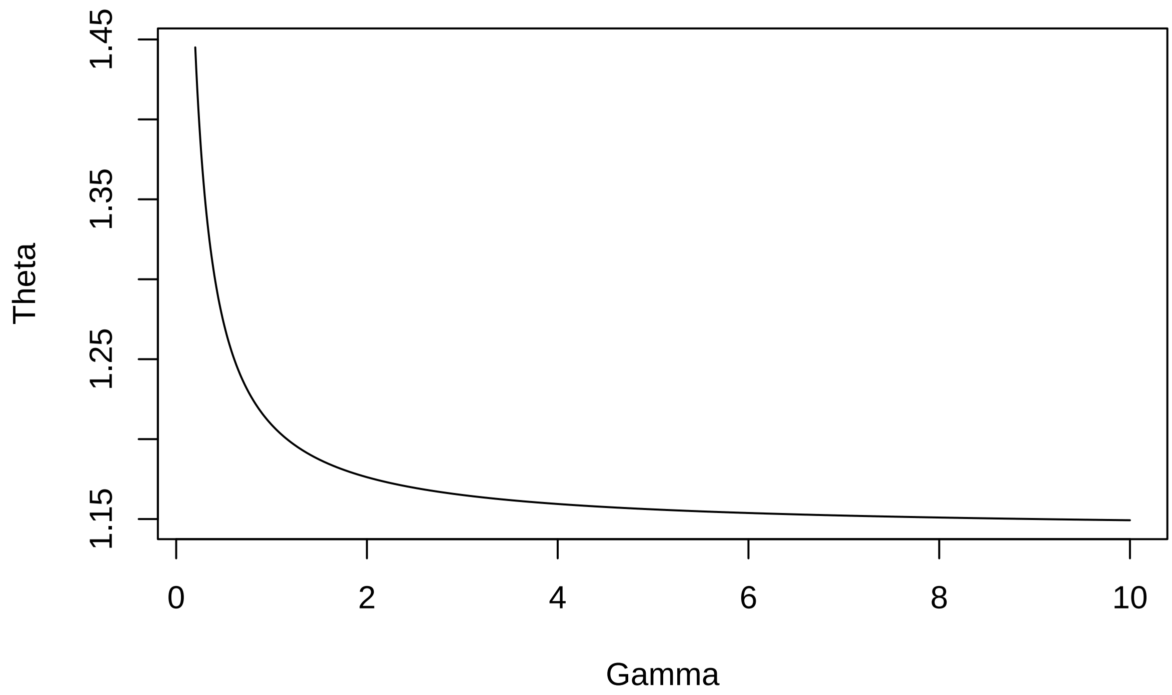



```
#Part One - Q3 section a
gamma <- c(0.1,2,100)
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)
P <- 1-pnbinom(ES+3*sqrt(VS),r,p)

#Part One - Q3 section b
gamma <- seq(2,10,by=0.01)
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)
P <- 1-pnbinom(ES+3*sqrt(VS),r,p)
```

```r
plot(gamma,P,type="l",xlab="Gamma",ylab="Probability")
abline(h=1-pnorm(3,0,1),col="red",lty=2)
```

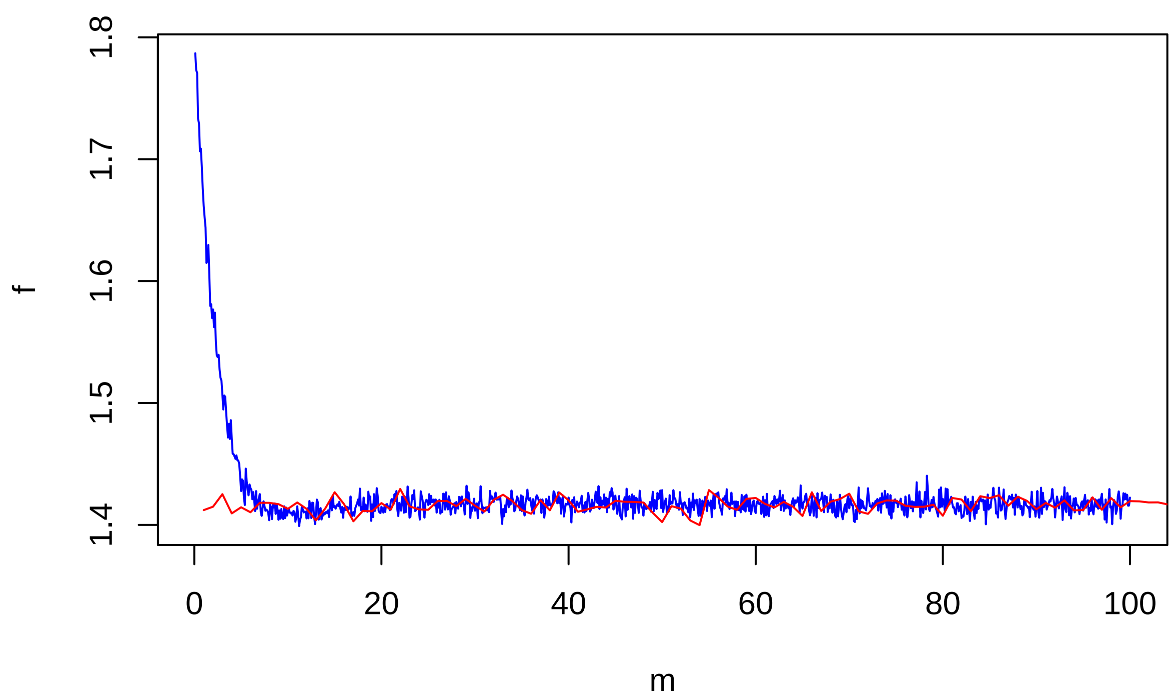



```r
#Part One - Q3 section c
gamma <- c(0.1,2,100)
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)
P <- 1-pnbinom(ES+3*sqrt(VS),r,p)

#Part One - Q4
gamma <- 0.5
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)

#Part One - Q5
gamma <- seq(0.2,10,by=0.01)
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)

theta <- qnorm(0.9,0,1)*sqrt(VS)/ES

plot(gamma,theta,type="l",xlab="Gamma",ylab="Theta")
```




```r
#Part One - Q6
gamma <- 0.5
ES <- (r*(1-p)*gamma)/(p*lambda)
VS <- (gamma*r*(1-p)*(gamma+p))/(p^2*lambda^2)
Skew <- (2*p^2+3*gamma*p+gamma^2*(2-p))/(sqrt(r*gamma*(1-p))*(p+gamma)^1.5)
alpha <- 4/Skew^2
beta <- sqrt(alpha/VS)
k <- ES-alpha/beta
theta <- qnorm(0.9,0,1)*sqrt(VS)/ES

#Part One - Q8
samplen <- 10000
gamma <- 0.5
mn <- 1000
unit <- 0.1
m <- seq(0,unit*mn,by=unit)
m <- m[-1]
f <- rep(0,mn)
g <- rep(0,mn)
Y <- rep(0,samplen)
Z <- rep(0,samplen)
S <- rep(0,samplen)
B <- 0.1
```

```r
mfound <- NULL
mSum <- 0

for(i in 1:mn){

  for(j in 1:samplen){
    Newsample <- sum(rgamma(rnbinom(1,r,p),gamma,lambda))
    Y[j] <- min(Newsample,m[i])
    Z[j] <- max(0,Newsample-m[i])
    S[j] <- Newsample
  }

  f[i] <- mean(exp(B*Y))*exp(2*B*mean(Z))
  g[i] <- mean(exp(B*S))

  if(f[i]>g[i]){
    mfound <- c(mfound,m[i])
  }
}
mSum <- mSum+max(mfound)+unit
plot(m,f, type = "l", col = "blue")
lines(g, col = "red")
```



```r
mfound
```

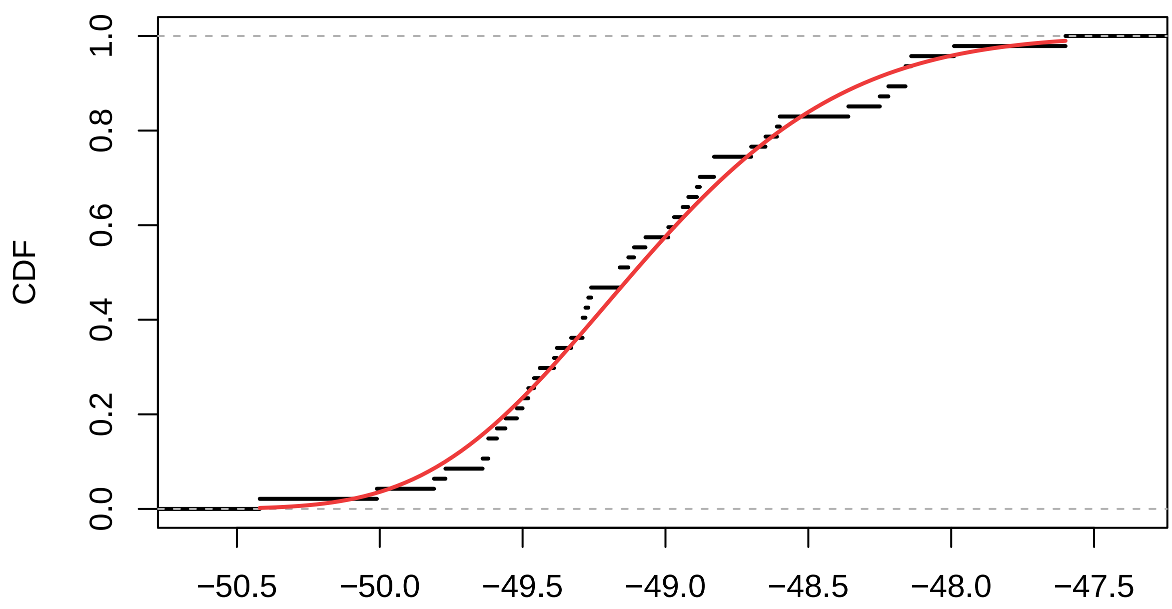
```
##  [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9
## [20] 2.0 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8
## [39] 3.9 4.0 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8 4.9 5.0 5.1 5.2 5.3 5.4 5.5 5.6 5.7
## [58] 5.8 5.9 6.0 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.9
```

```r
#Part Two - Q1: Please note that the codes in Questions 2.1, 2.2, 2.4 are heavily reliant on th
  dat <- read.csv("bestF400m.csv")
  dat$Perf <- -1*dat$Perf
  dat <- dat$Perf
  sample <- dat

  GEV.MM <- function(sample){
  # Empirical moments
  m1 <- mean(sample)
  m2 <- mean(sample^2)
  m3 <- mean(sample^3)
  var <- m2-m1^2
  skew <- mean((sample-m1)^3)/var^1.5
  # Declaring function g_k
  gk <- function(k, xi){gamma(1-k*xi)}
  # Function whose root we want to find
  func <- function(xi){
  sign(xi)*(gk(3,xi)-3*gk(2,xi)*gk(1,xi)+2*gk(1,xi)^3)/
  (gk(2,xi)-gk(1,xi)^2)^1.5 - skew
  }
  # Finding "xi"
  e.h <- uniroot(func, interval=c(-10, 1/3-0.00001))$root
  # Finding "sigma"
  s.h <- sqrt(var*e.h^2/(gk(2,e.h)-gk(1,e.h)^2))
  # Finding "mu"
  m.h <- m1 - s.h/e.h*(gk(1,e.h)-1)
  # Return results
  results <- list(m.h, s.h, e.h)
  names(results) <- c("mu", "sigma", "xi")
  return(results)
  }

  dGEV<- function(x,mu,sigma,xi){
    ifelse(1+xi*(x-mu)/sigma<=0,0,1/sigma*(1+xi*(x-mu)/sigma)^(-1/xi-1)*exp(-(1+xi*(x-mu)/sigma
  }

  nL <- function(p) {-prod(dGEV(dat, p[1],p[2],p[3]))}

  nll <- function(p) {-sum(log(dGEV(dat, p[1],p[2],p[3])))}
```

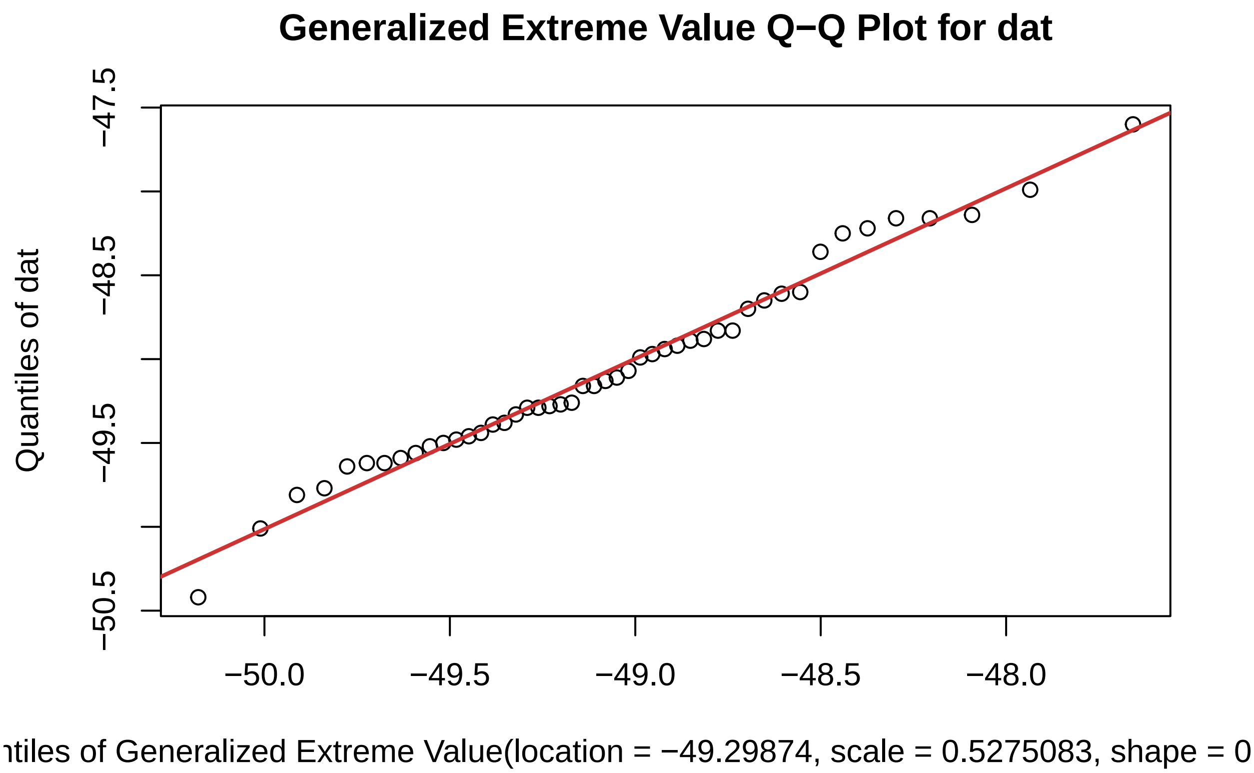
```r
  MM <- GEV.MM(dat)
  iv=c(MM$mu,MM$sigma,MM$xi)

  fit1 <- optim(iv,nL)
  fit2 <- optim(iv,nll)
```

```
## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced

## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced

## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced
```

```r
#Part Two - Q2
par <- iv
pFrechet <- function(x,m,s,e){ifelse(1+e*(x-m)/s <=0, 0, exp(-(1+e*(x-m)/s)^(-1/e)))}
plot(ecdf(dat), do.points = FALSE, main = "", ylab="CDF", xlab="", lwd=2)
curve(pFrechet(x,m=par[1], s=par[2], e=par[3]), from=min(dat), to=max(dat), col = "brown2", lwd
```




```r
qqPlot(dat, distribution='gevd', param.list=list(location=par[1], scale=par[2], shape=-par[3]),
       add.line=T, line.lwd=2, line.col='brown3')
```




```
#Part Two - Q4
dat <- read.csv("bestF400m.csv")
dat <- filter(dat,Year>=1990)
dat$Perf <- -1*dat$Perf
dat <- dat$Perf

MM <- GEV.MM(dat)
iv=c(MM$mu,MM$sigma,MM$xi)
fit1 <- optim(iv,nL)
fit2 <- optim(iv,nll)
```

```
## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced

## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced

## Warning in log(dGEV(dat, p[1], p[2], p[3])): NaNs produced
```