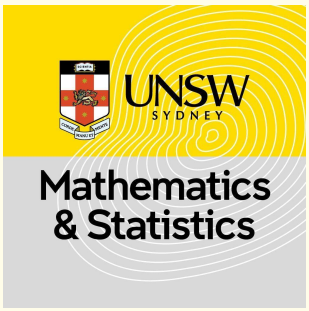


MATH5805
A Modern Introduction to Extreme Value Theory
Term 1, 2025

Dr. Boris Béranger

Week 7 - Lab Solutions

Introductory note
These multivariate extremes labs build on the use of the `R` package `extRemes` which focuses on univariate extremes, and use the packages `evd`, `mev`, `ExtremalDep` and `SpatialExtremes`. Each package has extensive `R` help files with a manual with many worked examples.

Exercise 1 - Standardisation

Some extremes packages in `R` will estimate marginal parameters simultaneously with dependence parameters; other packages will require that the margins are transformed to specific distributions and only fit the dependence parameters. A benefit of the latter is modelling and coding simplicity. A benefit of the former is statistical efficiency and modelling flexibility. Transforming margins leaves the dependence structure intact (think copulas).

If data X come from a distribution function F and we wish them to be transformed to a distribution function G, then the transformation is $G^{-1}(F(X))$.

- `mtransform` (`evd` package) transforms from GEV to exponential margins.
- `trans` (`extRemes` package) transforms to Gumbel or exponential margins.
- `spunif` (`mev` package) transforms to uniform margins.
- `trans2UFrechet` (`extremalDep` package) transforms GEV to unit Fréchet margins (the reverse is done throught `trans2GEV`).

All of the above functions assume that the GEV/GPD/PP limit for the data is valid, and transform the marginal data accordingly. Obviously transforming from (say) GEV to unit Fréchet when the data are not GEV-distributed, may not be the best option.

a) Write a function `x2u` that transforms the margins of samples from a multivariate distribution to uniform distributions.

Hint: use `rank()/(length()+1)`.

Solution:

```
x2u <- function(x){
 lx <- nrow(x)
 p <- ncol(x)
```

```
 u <- matrix(NA, nrow=lx, ncol=p)
 for(i in 1:p){
   u[,i] <- rank(x[,i], ties.method="random", na.last="keep")/(lx+1)
 }
 u
}
```

b) Write a function `u2f` that transforms the margins of samples from a multivariate distribution with uniform margins to unit Fréchet distributions ($GEV(1,1,1)$).

Solution:

```
u2f <- function(u){
  lu <- nrow(u)
  p <- ncol(u)
  f <- matrix(NA, nrow=lu, ncol=p)
  for(i in 1:p){
    f[,i] <- -1/log(u[,i])
  }
  f
}
```

c) Write a wrapper of the previous two functions `x2f` that transforms the margins of samples from a multivariate distribution to unit Fréchet margins.

Solution:

```
x2f <- function(x){
  u2f(x2u(x))
}
```

d) Draw $1,000$ samples from a bivariate normal distribution $(X_1, X_2) \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$, where $\mathbf{0}$ is the zero vector and $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, for some large value of ρ. Construct a bivariate scatterplot of the samples, plus histograms of both univariate margins. Transform to uniform margins and again produce the scatterplot and histograms. Transform from uniform to Fréchet margins and again produce the same plots (perhaps on the log scale). Fit a GEV distribution to the univariate Fréchet margins and show that they are Fréchet distributed.

Hint: use the `rmvnorm()` function from the `mvtnorm` package.

Solution:

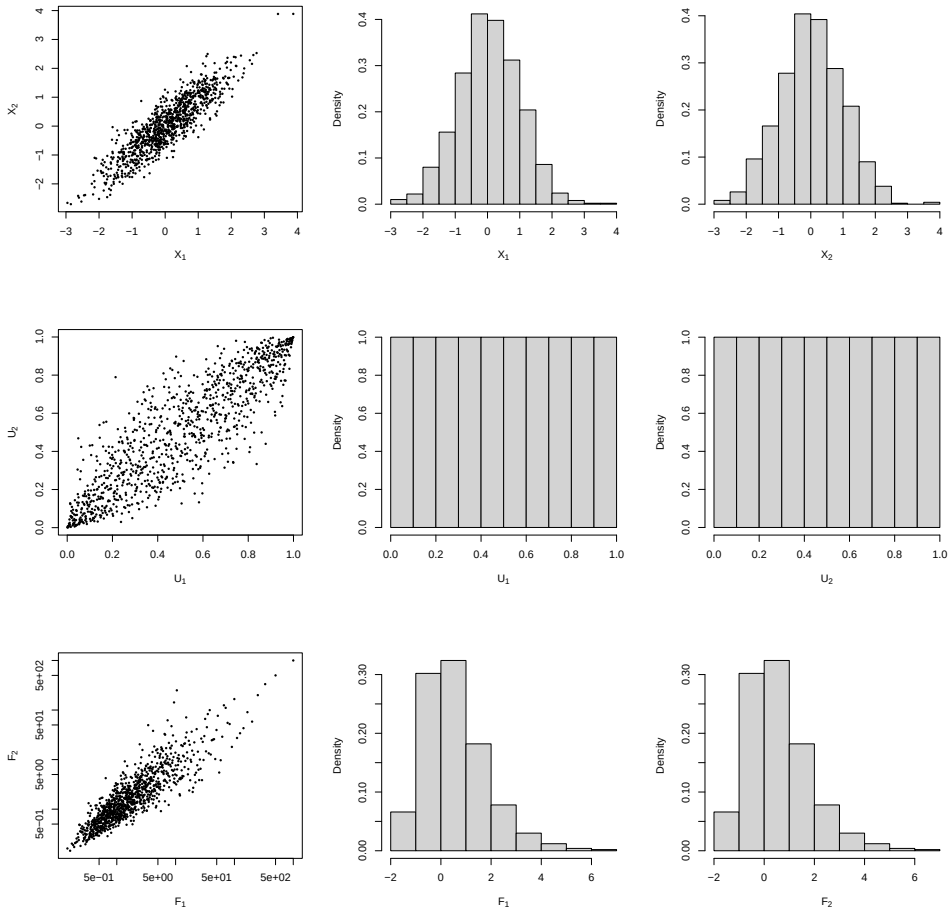
```
library(extRemes)
library(mvtnorm)
par(mfrow=c(3,3))

# Generate data
```

```r
set.seed(123)
rho <- 0.9
x <- rmvnorm(1000, mean=rep(0,2), sigma=matrix(c(1,rho,rho,1),ncol=2))
plot(x, cex=0.5, pch=16, xlab=expression(X[1]), ylab=expression(X[2]))
hist(x[,1], xlab=expression(X[1]), probability=T, main="")
hist(x[,2], xlab=expression(X[2]), probability=T, main="")

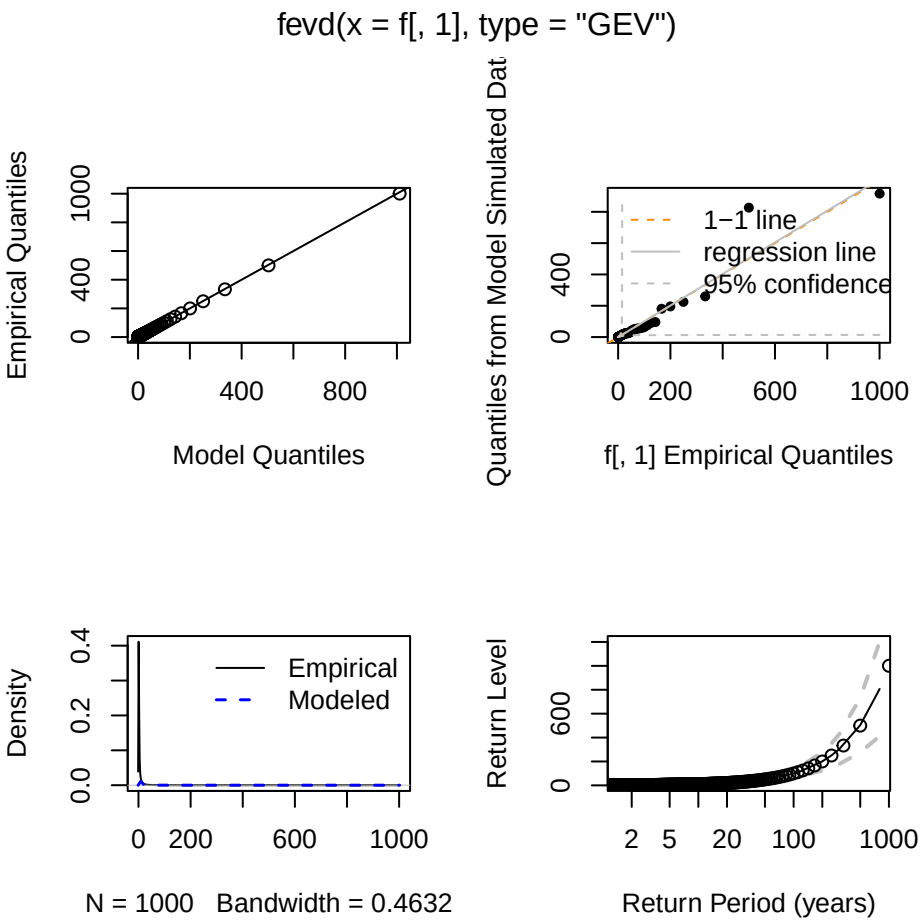
# Transform to uniform margins
u <- x2u(x)
plot(u, cex=0.5, pch=16, xlab=expression(U[1]), ylab=expression(U[2]))
hist(u[,1], xlab=expression(U[1]), probability=T, main="")
hist(u[,2], xlab=expression(U[2]), probability=T, main="")

# Transform to unit Frechet margins
f <- u2f(u)
plot(f, cex=0.5, pch=16, xlab=expression(F[1]), ylab=expression(F[2]),
     log="xy")
hist(log(f[,1]), xlab=expression(F[1]), probability=T, main="")
hist(log(f[,2]), xlab=expression(F[2]), probability=T, main="")
```

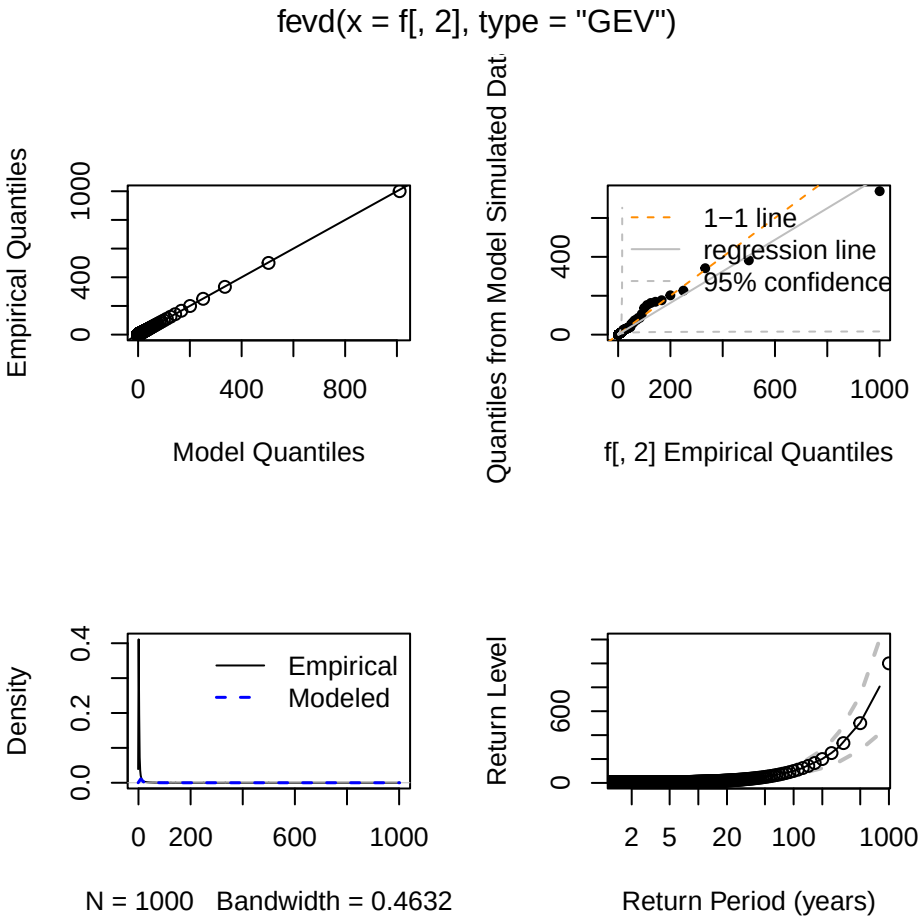



```r
# Fit GEV
fit1 <- fevd(f[,1], type="GEV")
```

```
plot(fit1)
```




```
fit2 <- fevd(f[,2],type="GEV")
plot(fit2)
```



Exercise 2 - Modelling maxima

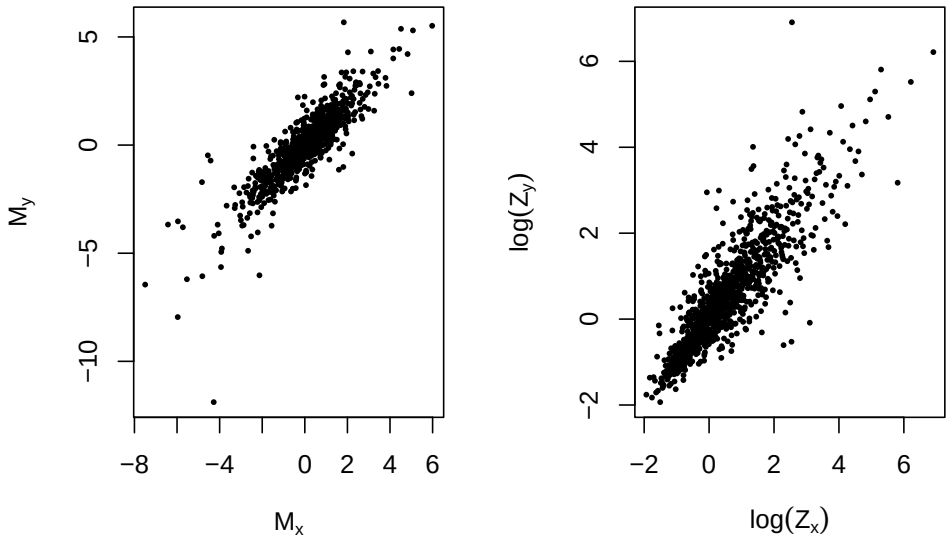
Suppose that $(M_{x,i}, M_{y,i})$, $i = 1, \ldots, 1000$, are a sequence of bivariate block maxima, where M_{xi} and M_{yi} are the margin-wise block maximum of n iid samples $(X, Y) \sim mt_\nu(\delta, \Sigma)$, where $mt_\nu(\delta, \Sigma)$ is a multivariate t-distribution with ν degrees of freedom and centrality and scale parameters δ and Σ. Suppose that $\delta = \mathbf{0}$ is the zero vector, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, for some large value of ρ, and $\nu = 3$.

a) For $n = 1$ generate 1000 bivariate block maxima and transform the margins to unit Fréchet scale.

Solution:

```
n <- 1 # block size
L <- 1000 # number of samples
rho <- 0.9
set.seed(123)
temp <- rmvt(L*n, sigma=matrix(c(1,rho,rho,1),ncol=2), df=3)
x <- matrix(temp[,1], ncol=n, byrow=T)
y <- matrix(temp[,2], ncol=n, byrow=T)
dat <- cbind( apply(x,1,max), apply(y,1,max))
```

```
par(mfrow=c(1,2))
plot(dat, pch=16, cex=0.5, ylab=expression(M[y]), xlab=expression(M[x]))
dat <- x2f(dat) # Transform to unit frechet margins
plot(log(dat), pch=16, cex=0.5, ylab=expression(log(Z[y])),
     xlab=expression(log(Z[x])))
```



b) Block maxima of t-distributed random variables are known to be in the domain of attraction of the extremal-t distribution. Fit the extremal-t distribution to the block maxima you generated and superimpose the fitted density over a scatterplot of the data. You may wish to plot on the log-scale. What are the fitted parameter values?

Hint: use the `fExtDep` function in the `ExtremalDep` package.

Solution:

```
library(ExtremalDep)
fit <- fExtDep(method="Composite", data=dat, model="ET", par.start=c(0,3), trace=2)
```

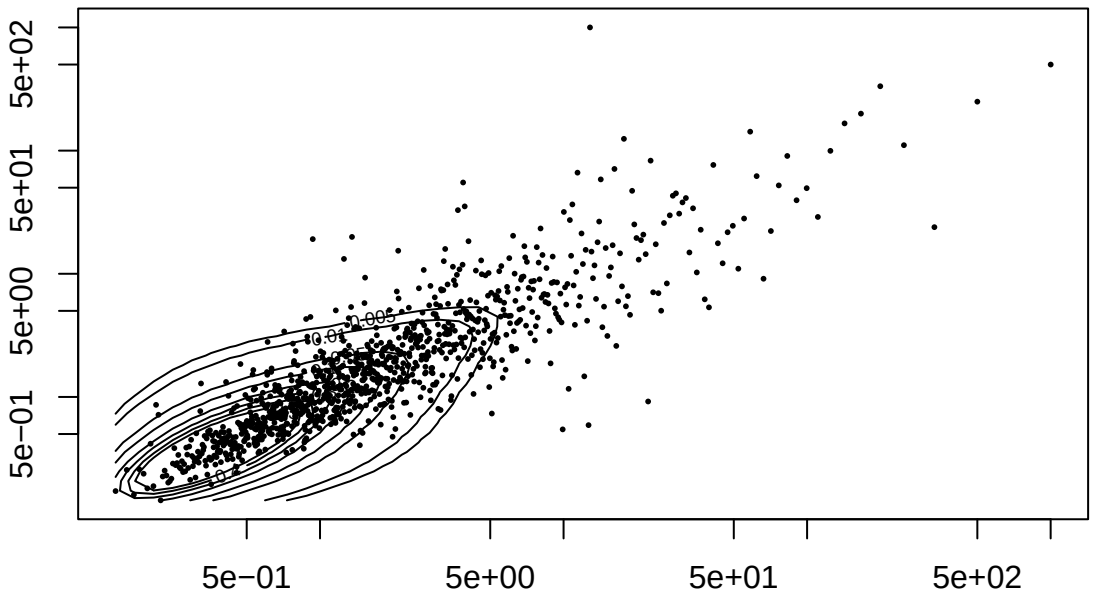
```
## initial  value 4152.493574
## iter  10 value 3546.096820
## final  value 3545.271708
## converged
```

```
fit$par
```

```
## [1] 0.937 2.450
```

```
xx <- exp( seq(log(min(dat[,1])), log(max(dat[,1])), length=50) )
yy <- exp( seq(log(min(dat[,2])), log(max(dat[,2])), length=50) )
aa <- expand.grid(xx,yy)
dd <- matrix(apply(aa, 1,dExtDep, model="ET", par=fit$par, angular=FALSE, log=FALSE),
             ncol=50,byrow=F)
# This plot doesn't seen quite right?
par(mfrow=c(1,1))
```

```
contour(xx, yy, dd, nlevels=10, levels=c(0.5,0.4,0.3,0.2,0.1,0.05,0.01,0.005), log="xy")
points(dat, cex=0.4, pch=16)
```

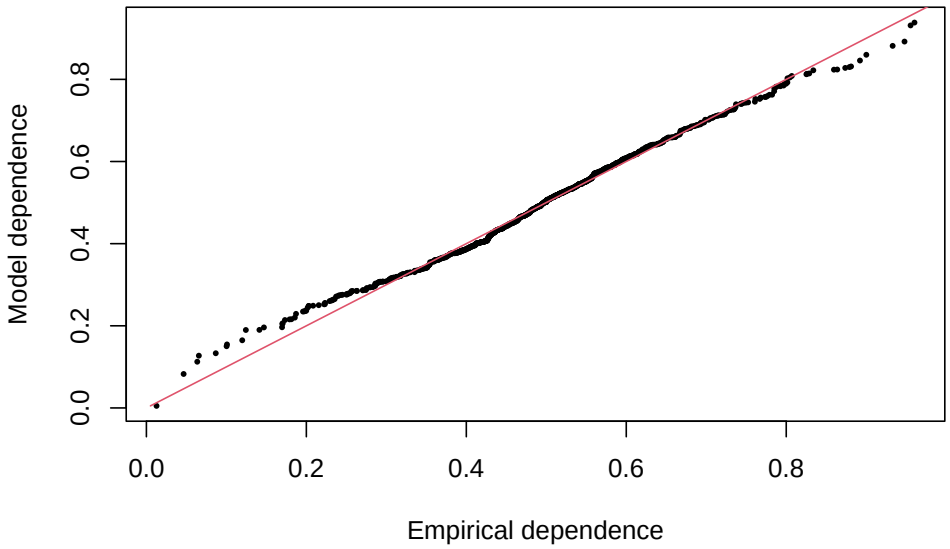


The fitted extremal-t density is broadly in the right location as the t-distributed data, though the centring does look a little off. The fitted parameter values are $\hat{\rho} = 0.940$, $\hat{\nu} = 2.594$ which is fairly close to the true parameters of $\rho = 0.9$ and $\nu = 3$. (This package does not produce confidence intervals.)

c) Generate 1000 samples $(S_{1,i}, S_{2,i})$, $i = 1, \ldots, 1000$, from the fitted extremal-t model (hint: use the function `rExtDep`). Compute the measure of extremal dependence represented by $W_i = S_{1i}/(S_{1i} + S_{2i})$. Compute the empirical version of this from the observed dataset and produce a qq-plot. Do the block maxima appear to be distributed according to the extremal-t limit?

Solution:

```
# ss <- rExtDep(n=L, model="ET", par=fit$par) # Slow!
# ww <- ss[,1]/(ss[,1]+ss[,2])
ww <- rExtDep(n=L, model="ET", par=fit$par, angular=TRUE)
wdat <- dat[,1]/(dat[,1]+dat[,2])
plot(sort(wdat), sort(ww[,1]), cex=0.5, pch=16, xlab="Empirical dependence",
     ylab="Model dependence")
r <- range(ww,wdat)
lines(r, r, col=2)
```



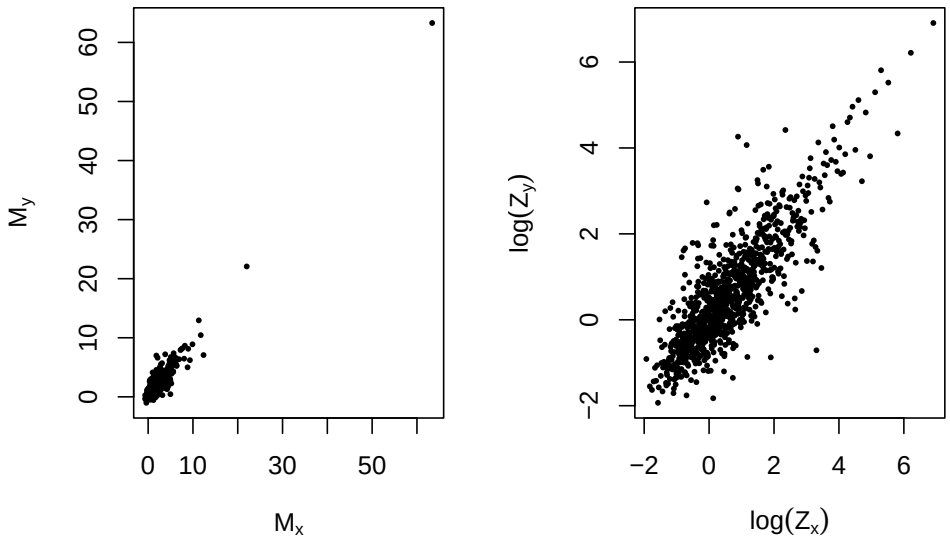
The fit (for this very specific measure) doesn't look too bad, though it is a little out in the tails.

d) Repeat steps a)–c) for increased block maxima sizes $n = 5, 20, 50 \ldots$ When do the block maxima converge to the extremal-t limit?

Note: as we are transforming the margins to unit Fréchet for each n, we are only examining convergence in the dependence structure. In practice, of course, we require both.

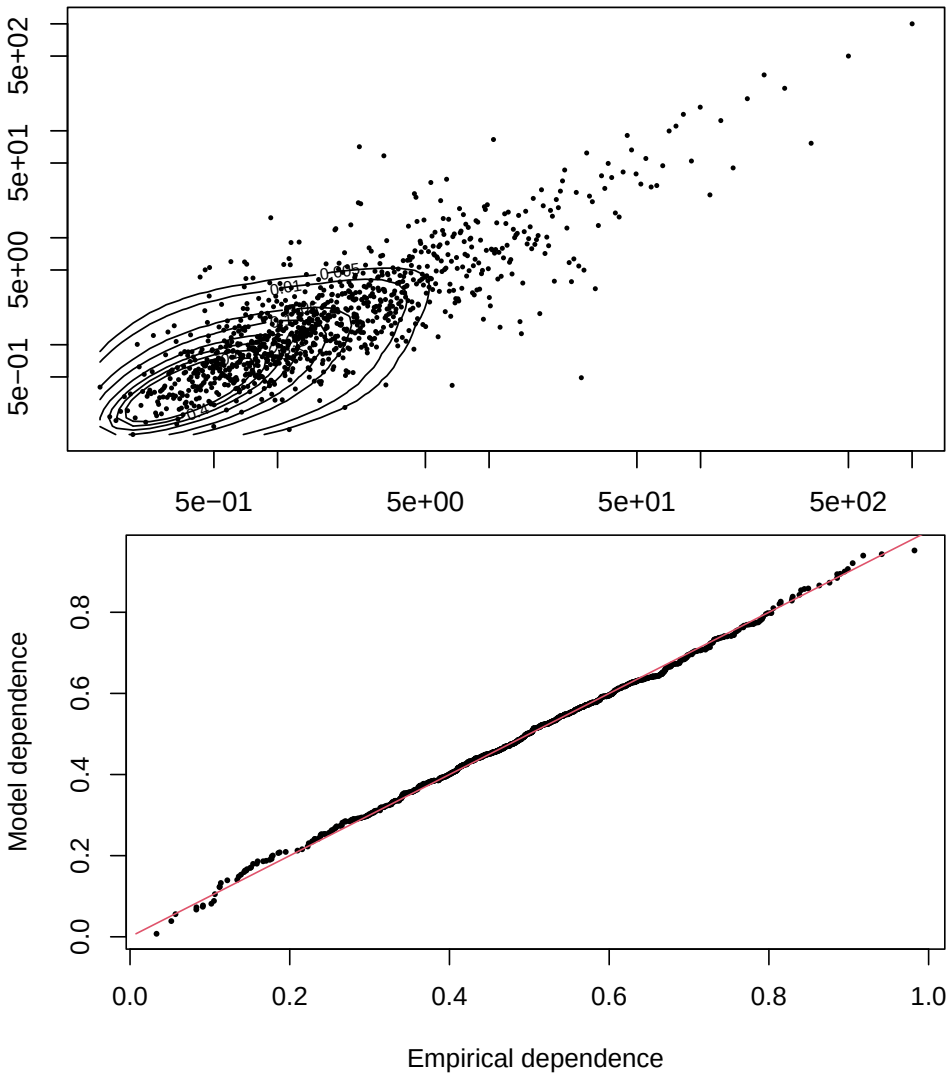
Solution:

For $n = 5$:



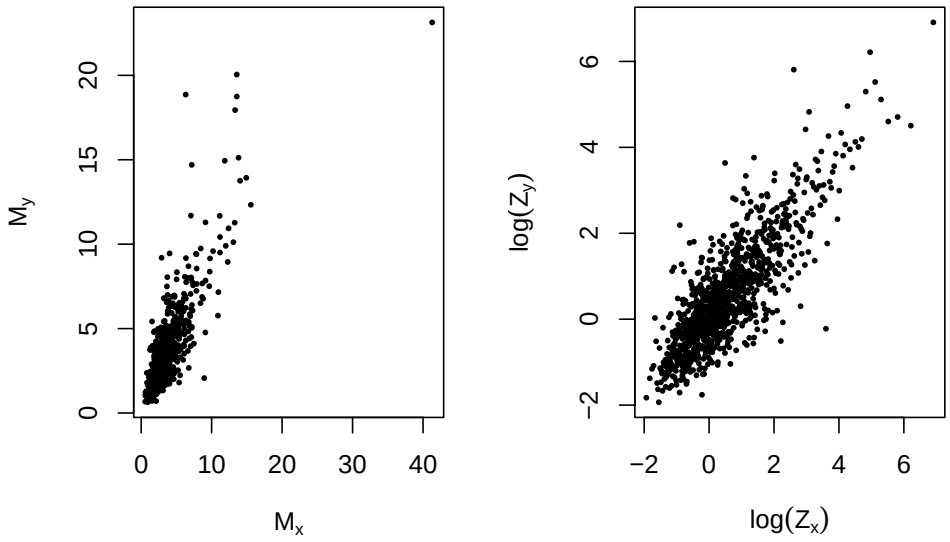

```
## initial  value 4155.925995
## iter  10 value 3674.768059
## final  value 3674.321291
## converged
```

```
## [1] 0.908 2.437
```



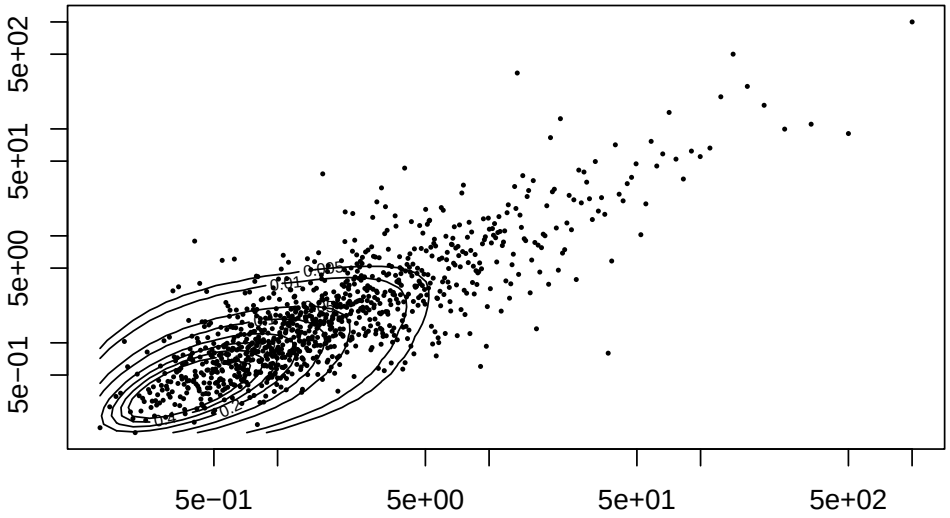
The fitted parameter values are $\hat{\rho} = 0.908$, $\hat{\nu} = 2.437$ which is relatively good but there is likely some variability. Also, it looks as though we have an improvemente in the tails of the qq-plot.

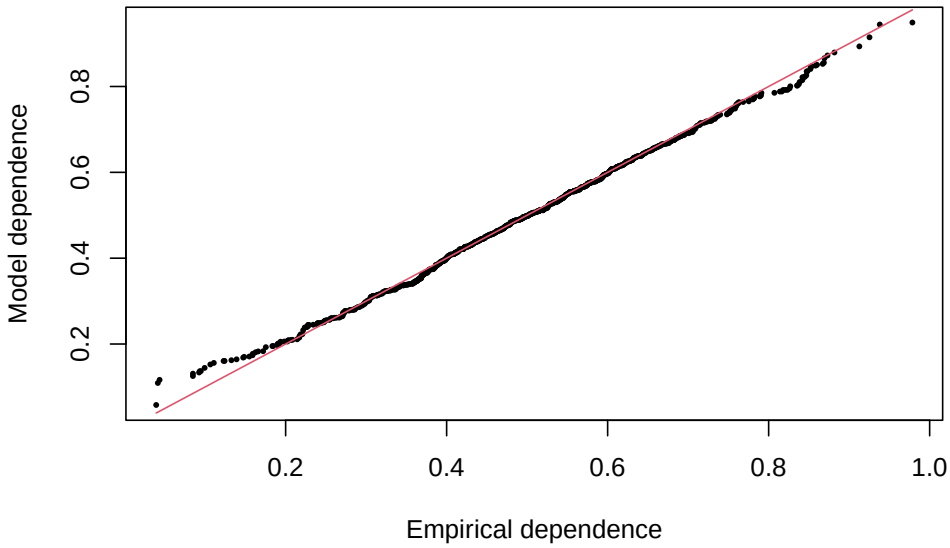
For $n = 20$:




```
## initial  value 4159.346729
## iter  10 value 3733.518920
## final  value 3732.441006
## converged
```

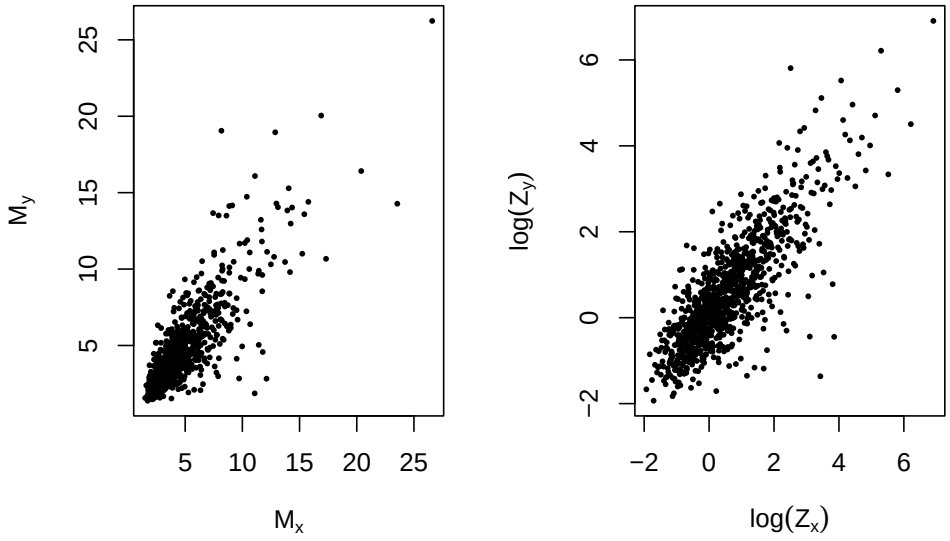
```
## [1] 0.922 3.678
```





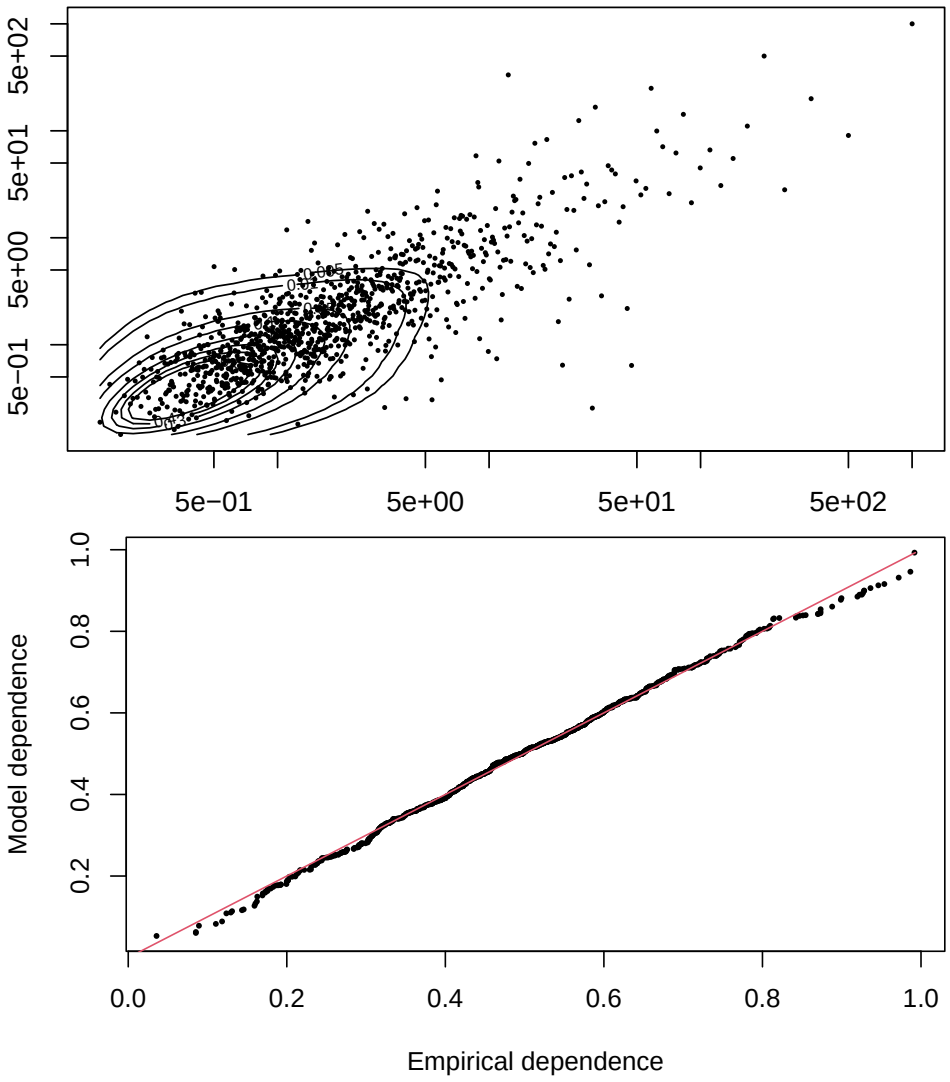
The fitted parameter values are $\hat{\rho} = 0.922$, $\hat{\nu} = 3.678$ which is fairly close to the true parameters of $\rho = 0.9$ and $\nu = 3$, but not as good as when $n = 5$. Again this is likely due to variability. Same comments about the qq-plot.

For $n = 50$:




```
## initial  value 4162.393703
## final  value 3744.056362
## converged
```

```
## [1] 0.901 2.780
```



The fitted parameter values are $\hat{\rho} = 0.901$, $\hat{\nu} = 2.780$ which is fairly close to the true parameters of $\rho = 0.9$ and $\nu = 3$. qq-plot is looking good. From this measure it looks as though the extremal limit has been achieved. The plot of fitted density over data points still looks centred in the wrong place, even as n increases. Is there a mistake?

Exercise 3 - Pickands dependence function of Asymmetric logistic model

Consider the bivariate ($d = 2$) asymmetric logistic model. The Pickands dependence function is given by

$$A(t) = (1 - \beta_1)t + (1 - \beta_2)(1 - t) + [(\beta_1 t)^\alpha + (\beta_2(1 - t))^\alpha]^{1/\alpha}, \quad 0 \leq t \leq 1$$

a) Plot the Pickands dependence function for various values of $\theta = (\beta_1, \beta_2, \alpha)$, $\alpha \geq 1$, $0 \leq \beta_1, \beta_2 \leq 1$.

Note: The `index.ExtDep` function from the `ExtremalDep` package can only plot the Pickands function for the Hüsler-Reiss, Extremal-t and Extremal skew-t models

Solution:

```
A.alog <- function(t,alpha,beta) {
  (1-beta[1])*t + (1-beta[2])*(1-t) +
    ((beta[1]*t)^alpha + (beta[2]*(1-t))^alpha)^(1/alpha)
}

par(mfrow=c(2,2))

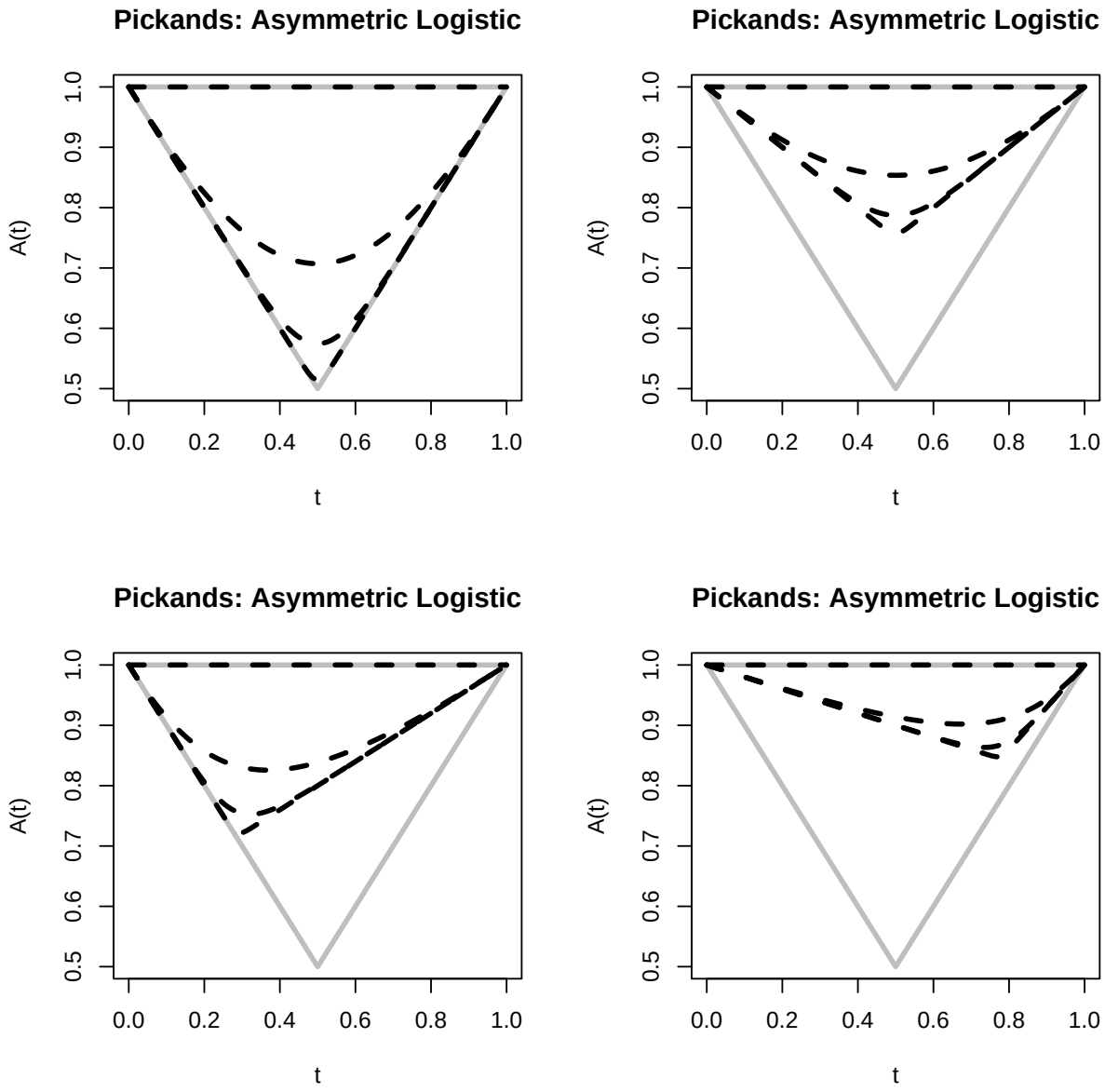
tt <- seq(0,1,length=100)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Asymmetric Logistic")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.alog(tt,alpha=1,beta=c(1,1)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=2,beta=c(1,1)), lty=2, lwd=3 )
lines(tt, A.alog(tt,alpha=5,beta=c(1,1)), lty=2, lwd=3 )
lines(tt, A.alog(tt,alpha=25,beta=c(1,1)), lty=2, lwd=3 )

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Asymmetric Logistic")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.alog(tt,alpha=1,beta=c(0.5,0.5)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=2,beta=c(0.5,0.5)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=5,beta=c(0.5,0.5)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=25,beta=c(0.5,0.5)), lty=2, lwd=3)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Asymmetric Logistic")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.alog(tt,alpha=1,beta=c(1,0.4)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=2,beta=c(1,0.4)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=5,beta=c(1,0.4)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=25,beta=c(1,0.4)), lty=2, lwd=3)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Asymmetric Logistic")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.alog(tt,alpha=1,beta=c(0.2,0.7)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=2,beta=c(0.2,0.7)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=5,beta=c(0.2,0.7)), lty=2, lwd=3)
lines(tt, A.alog(tt,alpha=25,beta=c(0.2,0.7)), lty=2, lwd=3)
```



b) What condition is needed to have a symmetric dependence structure? Can we achieve near/complete dependence/independence?

Solution:
Complete independence can be obtained for $\alpha = 1$ or $\beta_1 = 0$ or $\beta_2 = 0$. Complete dependence can be obtained as $\alpha \to \infty$ if $\beta_1 = \beta_2 = 0.5$. Symmetry is obtained if $\beta_1 = \beta_2$.

Exercise 4 - Pickands dependence function of Hüsler-Reiss model

Consider the bivariate ($d = 2$) Hüsler-Reiss model. The Pickands dependence function is given by

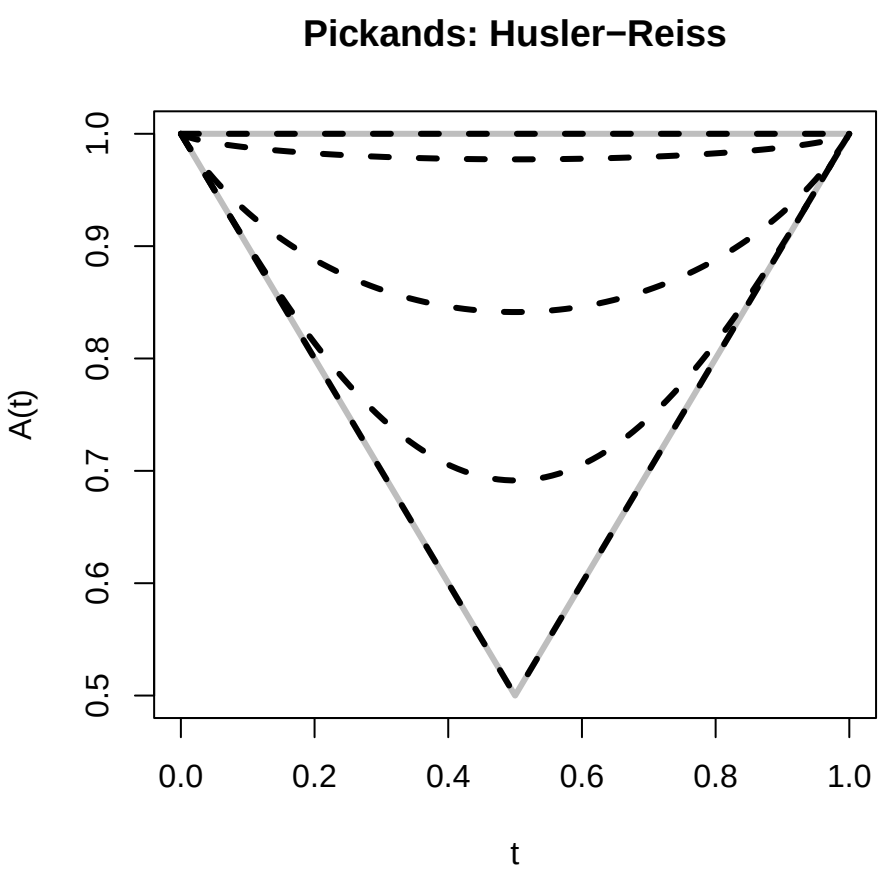
$$A(t) = t\Phi\left(\lambda + \frac{\log(t/(1-t))}{2\lambda}\right) + (1-t)\Phi\left(\lambda + \frac{\log((1-t)/t)}{2\lambda}\right), \quad 0 \le t \le 1$$

a) Plot the Pickands dependence function for various values of $\theta = \lambda$, $\lambda > 0$.

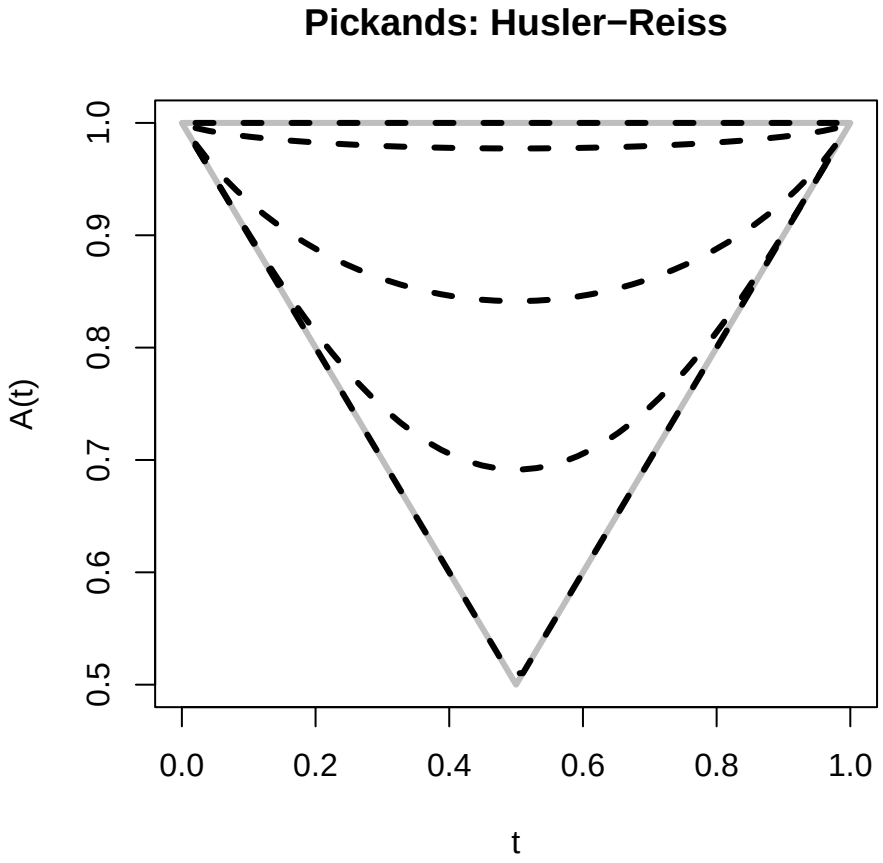
Solution:

```
A.hr <- function(t,lambda){
  t*pnorm(lambda+log(t/(1-t))/(2*lambda)) + (1-t)*pnorm(lambda+log((1-t)/t)/(2*lambda))
}
tt <- seq(0,1,length=100)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Husler-Reiss")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.hr(tt,lambda=0.001), lty=2, lwd=3)
lines(tt, A.hr(tt,lambda=0.5), lty=2, lwd=3)
lines(tt, A.hr(tt,lambda=1), lty=2, lwd=3)
lines(tt, A.hr(tt,lambda=2), lty=2, lwd=3)
lines(tt, A.hr(tt,lambda=5), lty=2, lwd=3)
```




```
# Alternative using ExtremalDep package
W <- simplex(d=2)
plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Husler-Reiss")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="HR", par=0.001), lty=2, lwd=3)
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="HR", par=0.5), lty=2, lwd=3)
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="HR", par=1), lty=2, lwd=3)
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="HR", par=2), lty=2, lwd=3)
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="HR", par=5), lty=2, lwd=3)
```



b) What condition is needed to have a symmetric dependence structure? Can we achieve near/complete dependence/independence?

Solution:
We can obtain complete dependence as $\lambda \to 0$ and complete dependence as $\lambda \to \infty$.

Exercise 5 - Pickands dependence function of Extremal-t model

Consider the bivariate ($d = 2$) Extremal-t model. The Pickands dependence function is given by

$$A(t) = tT_{\nu+1}\left[\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{t}{1-t}\right)^{1/\nu} - \rho\right)\right] + (1-t)T_{\nu+1}\left[\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{1-t}{t}\right)^{1/\nu} - \rho\right)\right], \quad 0 \le t \le 1,$$

where T_ν is the standard (univariate) t distribution function.

a) Plot the Pickands dependence function for various values of $\theta = (\rho, \nu)$, $\rho \in [-1, 1]$, $\nu > 0$.

Solution:

```
A.et <- function(t,rho,nu){
  s <- sqrt((nu+1)/(1-rho^2))
  t*pt(s*((t/(1-t))^(1/nu)-rho),df=nu+1) + (1-t)*pt(s*(((1-t)/t)^(1/nu)-rho),df=nu+1)
}

par(mfrow=c(2,2))
```

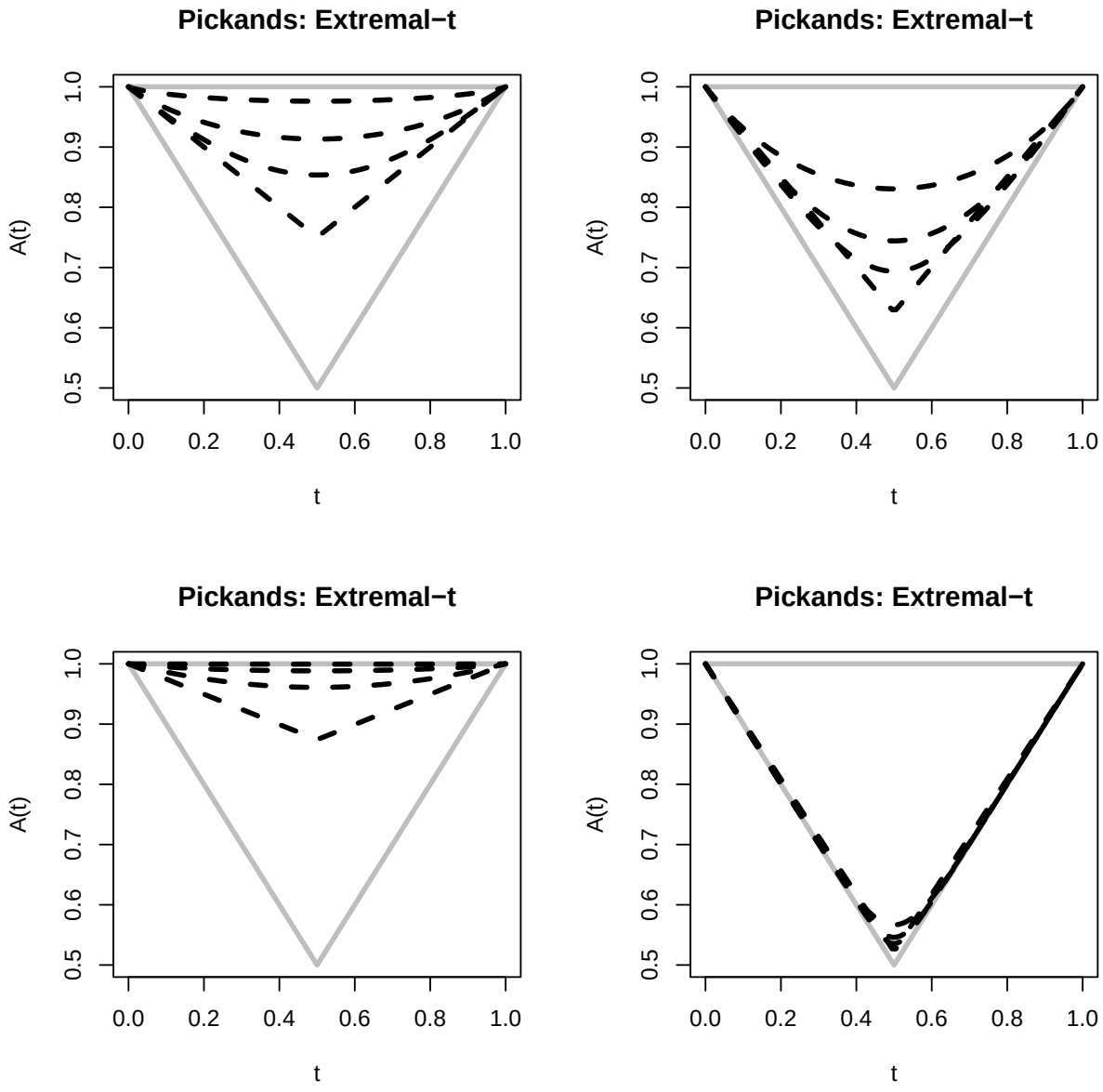
```r
tt <- seq(0,1,length=100)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Extremal-t")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.et(tt,rho=0,nu=.00001), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0,nu=1.0), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0,nu=2.1), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0,nu=5.1), lty=2, lwd=3)

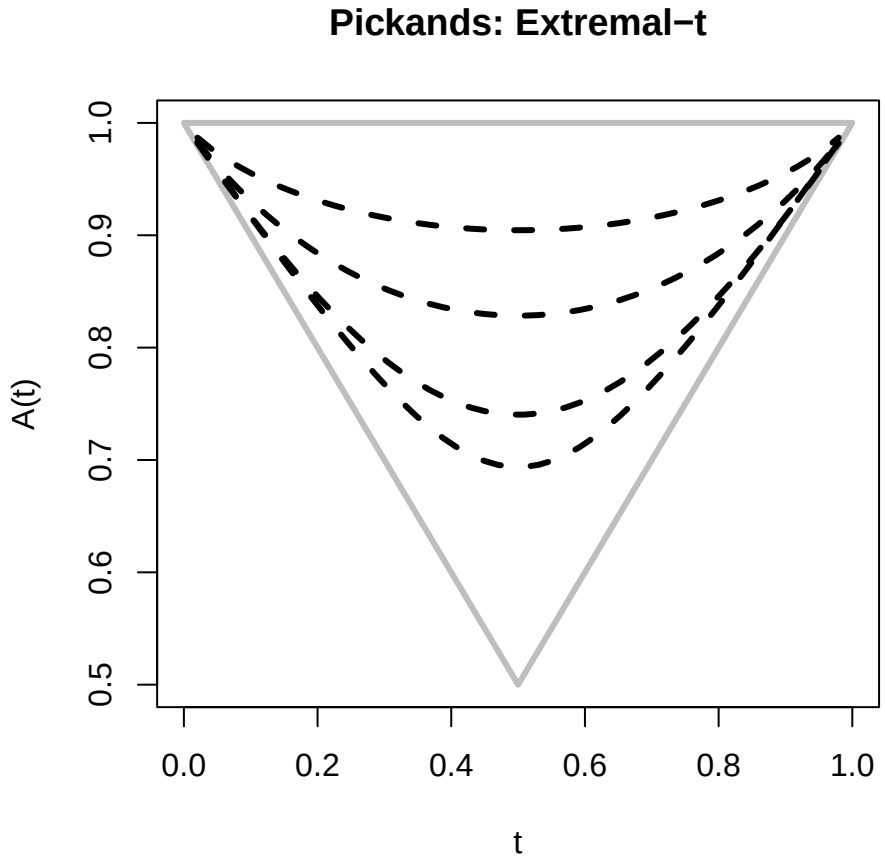
plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Extremal-t")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.et(tt,rho=0.7,nu=.01), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0.7,nu=1.0), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0.7,nu=2.1), lty=2, lwd=3)
lines(tt, A.et(tt,rho=0.7,nu=5.1), lty=2, lwd=3)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Extremal-t")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt, A.et(tt,rho=-0.7,nu=.01), lty=2, lwd=3)
lines(tt, A.et(tt,rho=-0.7,nu=1.0), lty=2, lwd=3)
lines(tt, A.et(tt,rho=-0.7,nu=2.1), lty=2, lwd=3)
lines(tt, A.et(tt,rho=-0.7,nu=5.1), lty=2, lwd=3)

plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Extremal-t")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(tt,A.et(tt,rho=0.99,nu=.01), lty=2, lwd=3)
lines(tt,A.et(tt,rho=0.99,nu=1.0), lty=2, lwd=3)
lines(tt,A.et(tt,rho=0.99,nu=2.1), lty=2, lwd=3)
lines(tt,A.et(tt,rho=0.99,nu=5.1), lty=2, lwd=3)
```




```
# Alternative using ExtremalDep package
W <- simplex(d=2)
plot(NA, NA, xlim=c(0,1), ylim=c(0.5, 1), ylab="A(t)", xlab="t",
     main="Pickands: Extremal-t")
polygon(c(0, 0.5, 1), c(1, 0.5, 1), lwd=3, border = 'grey')
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="ET", par=c(0.7, 1)), lty=2, lwd=3
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="ET", par=c(0.7, 2)), lty=2, lwd=3
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="ET", par=c(0.7, 5)), lty=2, lwd=3
lines(W[,1], apply(W, 1, index.ExtDep, object="pickands", model="ET", par=c(0.7, 10)), lty=2, lwd=
```



b) Can we achieve near/complete dependence/independence?

Solution:
From the above it can be seen that dependence increases as ρ increases from -1 to 1 (there is no symmetry in ρ here). In addition, dependence increases as $\nu \to 0$. Hence, we can get complete independence as $\nu \to \infty$ and(?) $\rho \to -1$, and complete dependence as $\nu \to 0$ and $\rho \to 1$. Note that for a d-dimensional Pickands' function, the ρ parameter controls dependence between a given pair of dimensions (ρ_{ij}) whereas ν provides an overall level of d-dimensional dependence.