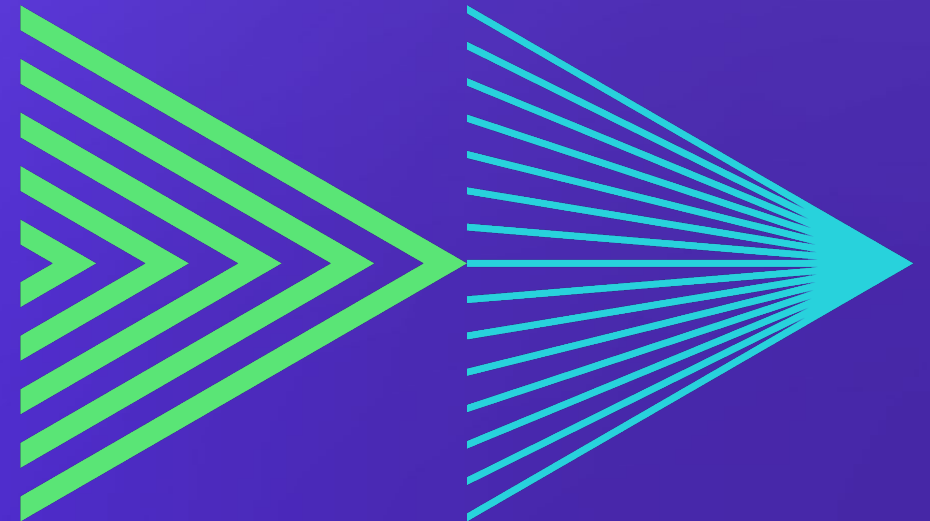


어떤 서비스를 선택해야 이용자 감소 리스크가 최소가 될까? (Feat. VAR-Copula 모형)

NHN PAYCO 금융사업팀
윤덕주



1. Introduction
2. VAR 모형
3. Copula함수
4. 모델링 설계
5. R프로그램을 이용한 서비스 간 이용자 감소 리스크(VAR-Copula) 측정 및 해석

앱 서비스 트렌드

=

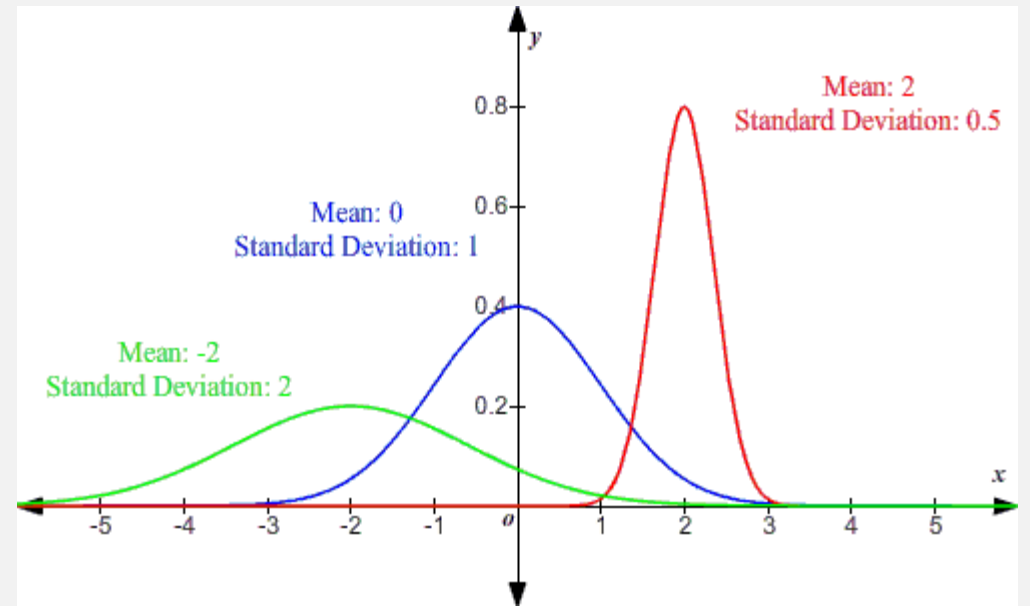
**여러 가지 서비스를 통합하여
하나의 앱으로 출시**

**출시한 다양한 서비스 중
일부만 운영해야 한다면?
어떤 서비스를 선택해야
이용 감소 리스크가 최소가 될까?**

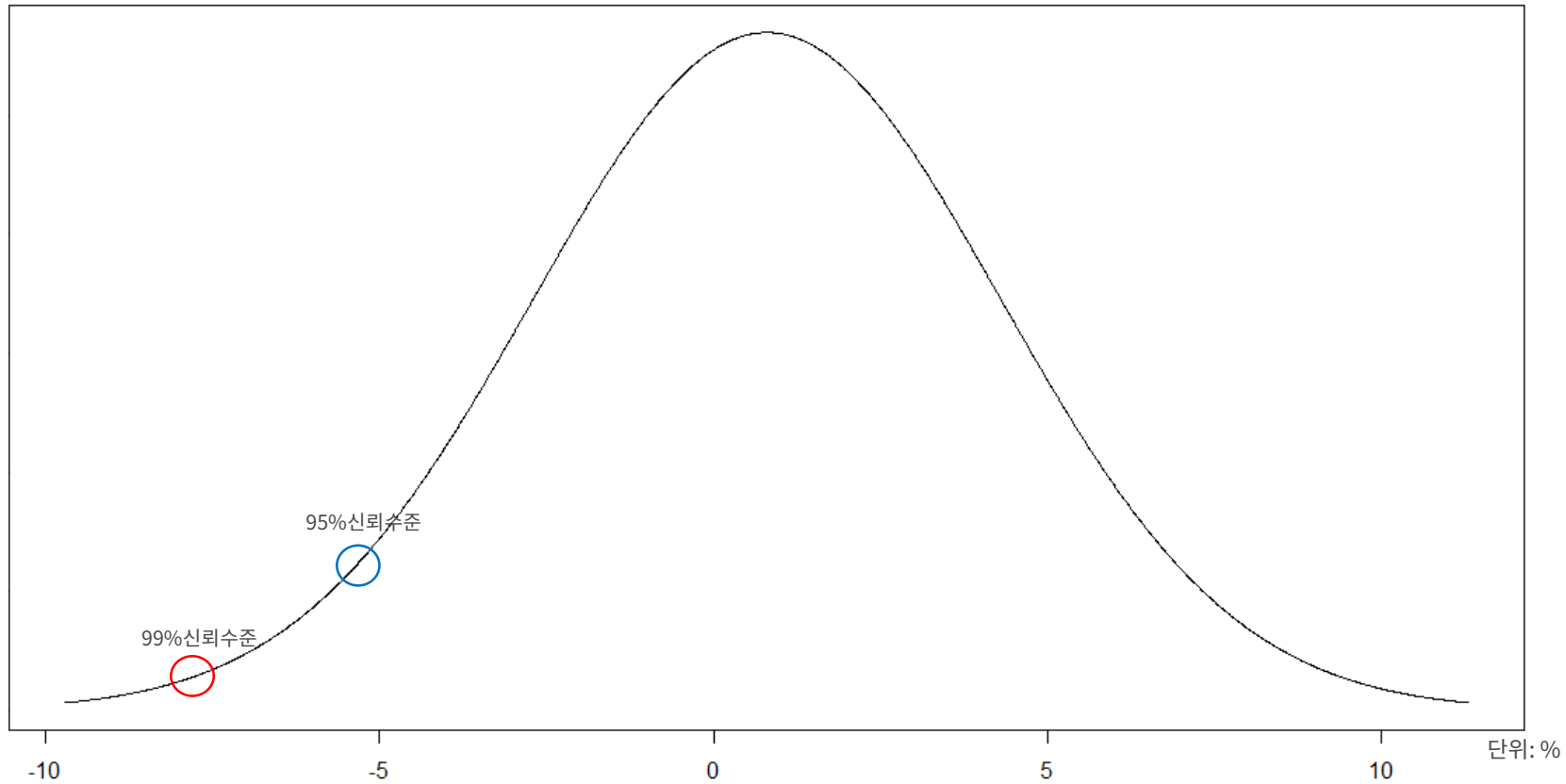
생존 분석?

머신 러닝?

정상적인 시장 여건 아래에서 주어진
신뢰 수준에서 일정 기간 발생할 수 있는
최대 손실 금액



삼성전자 최근 3개월 일별 수익률 정규 분포 가정



3개월간 1억을 삼성전자에 투자하면 1% VAR는 8,216,584원

삼성전자 5천만 원, 신한은행 5천만 원
분산 투자하면
1% VAR는 5,926,084원

*상관계수 0.4281

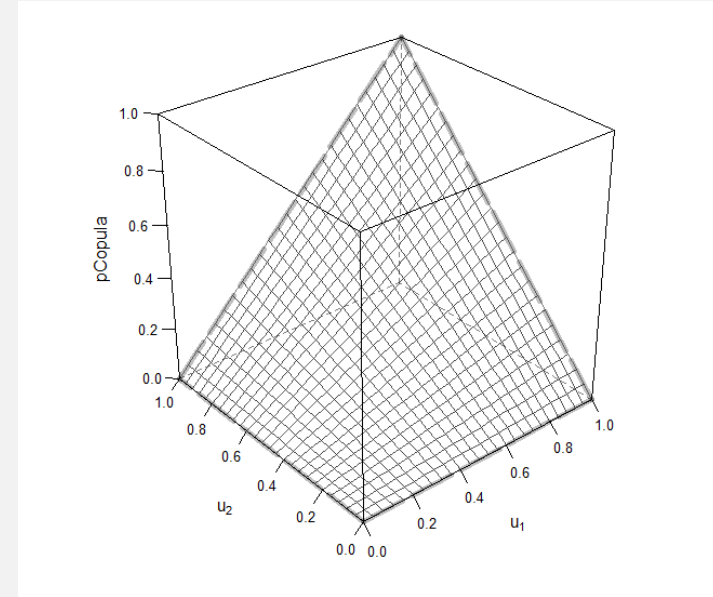
*1일 포트폴리오 $VAR_p = \sqrt{VAR_a^2 + VAR_b^2 + 2*pab*VAR_a*VAR_b}$

Copula 함수

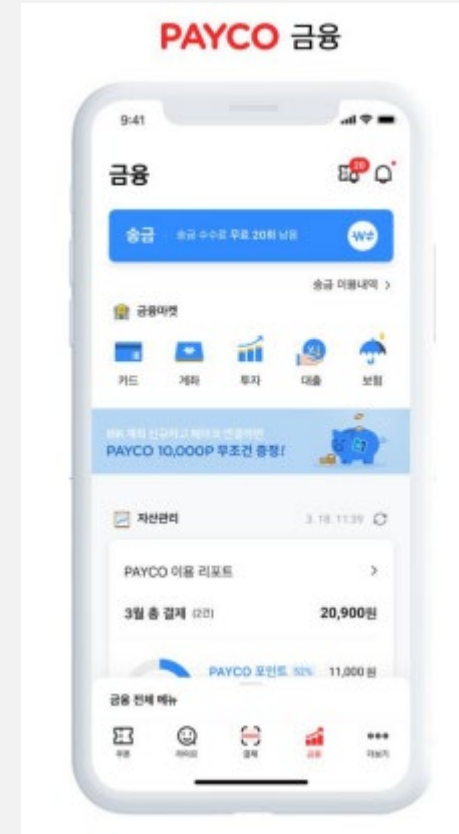
개별 주변부 확률 분포와 다변량 결합 분포를
결합해 주는 함수

*Skarl의 정리:

모든 $x, y \in \mathbb{R}$ 에 대하여 $F(x, y) = C(F_1(x), F_2(y))$ 를 만족
하는 C 가 존재하며 F_1, F_2 가 연속이면 C 는 유일하게 결정

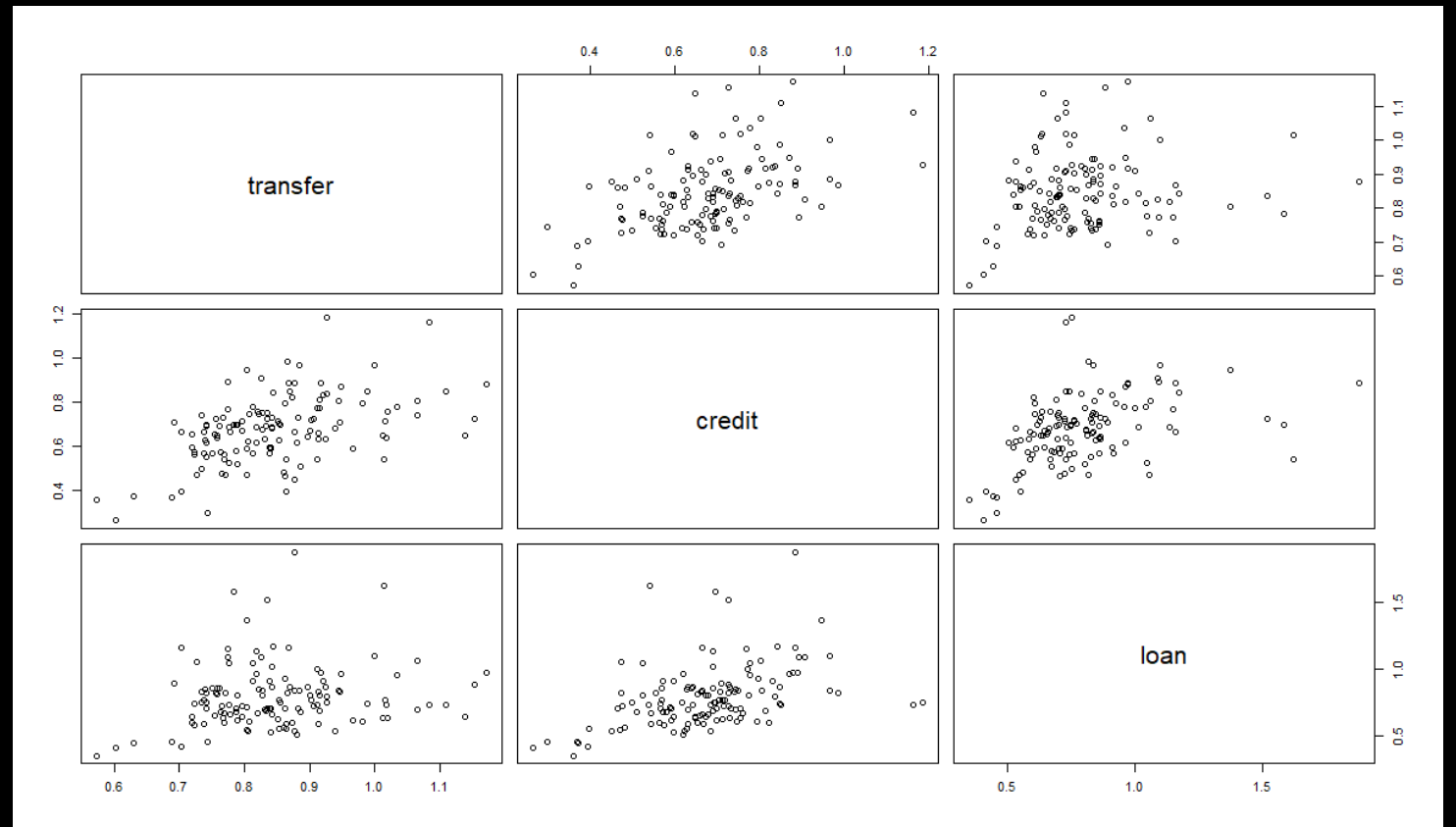


1. 서비스 이용률의 정의
2. 측정 서비스 선택
3. 데이터 가공
4. 알고리즘 적용
5. 해석



EDA 및 상관계수 추정

```
library(psych)  
pairs(payco)
```



EDA 및 상관계수 추정

```
summary(payco)
p<-corr.test(payco, use= "complete",
method="pearson",adjust="none")

s<-corr.test(payco, use= "complete",
method="spearman",adjust="none")

k<-corr.test(payco, use= "complete",
method="kendall", adjust="none")

p$r
s$r
k$r
```

```
transfer      credit      loan
Min.   :0.5727  Min.   :0.2635  Min.   :0.3520
1st Qu.:0.7689  1st Qu.:0.5860  1st Qu.:0.6480
Median :0.8365  Median :0.6816  Median :0.7486
Mean   :0.8460  Mean   :0.6763  Mean   :0.7946
3rd Qu.:0.9074  3rd Qu.:0.7545  3rd Qu.:0.8659
Max.   :1.1708  Max.   :1.1862  Max.   :1.8827
```

```
      transfer  credit  loan
transfer 1.0000000 0.4937495 0.1651402
credit   0.4937495 1.0000000 0.4207544
loan     0.1651402 0.4207544 1.0000000
```

```
      transfer  credit  loan
transfer 1.0000000 0.4708762 0.1517201
credit   0.4708762 1.0000000 0.4627801
loan     0.1517201 0.4627801 1.0000000
```

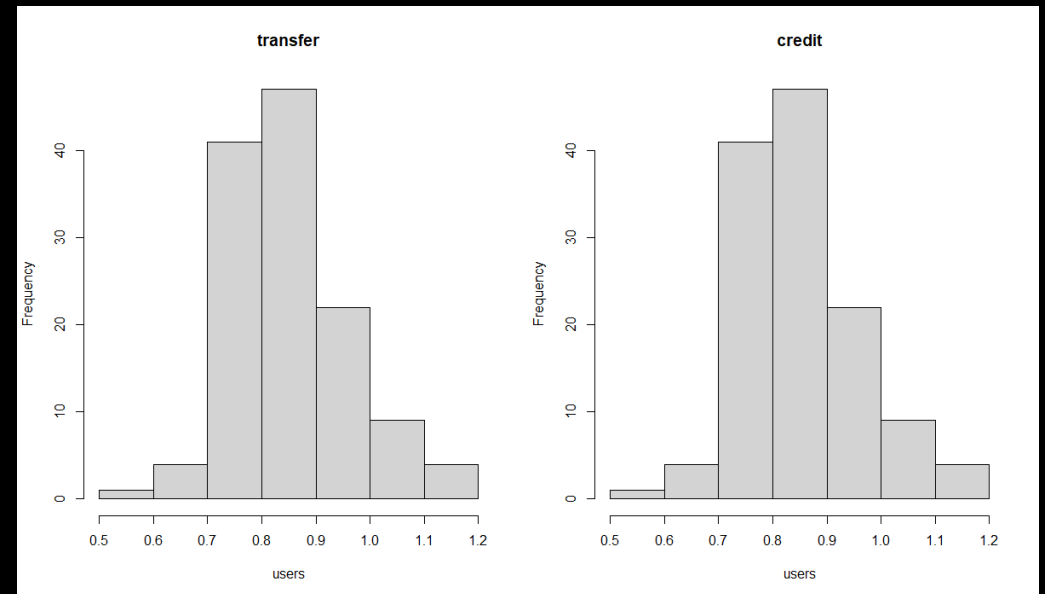
```
      transfer  credit  loan
transfer 1.0000000 0.3275682 0.1008348
credit   0.3275682 1.0000000 0.3281873
loan     0.1008348 0.3281873 1.0000000
```

Transfer-Credit / 히스토그램, 정규성 검증

```
library(quantmod)
x <- as.vector(payco$transfer)
y <- as.vector(payco$credit)

par(mfrow= c(1,2))
hist(x, main= "transfer", xlab= "users")
hist(x, main= "credit", xlab= "users")

shapiro.test(x)
shapiro.test(y)
```



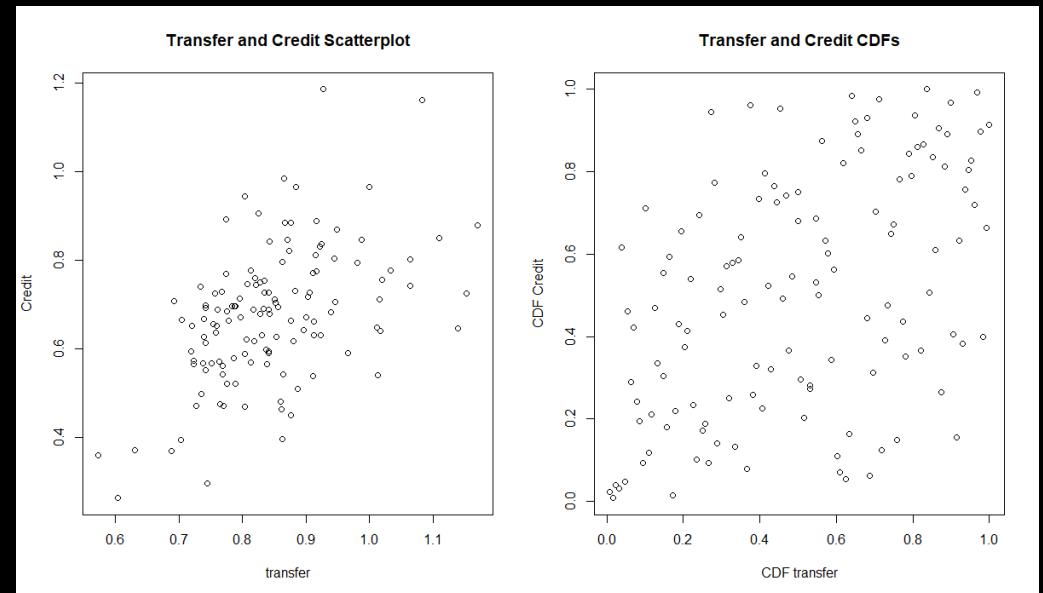
```
data: x
W = 0.96607, p-value = 0.002691
```

```
data: y
W = 0.98204, p-value = 0.08801
```

Transfer-Credit / correlation structure

```
par(mfrow=c(1,2))  
plot(x,y,xlab="transfer",ylab="transportation",  
main="transfer and transfortation Scatterplot")
```

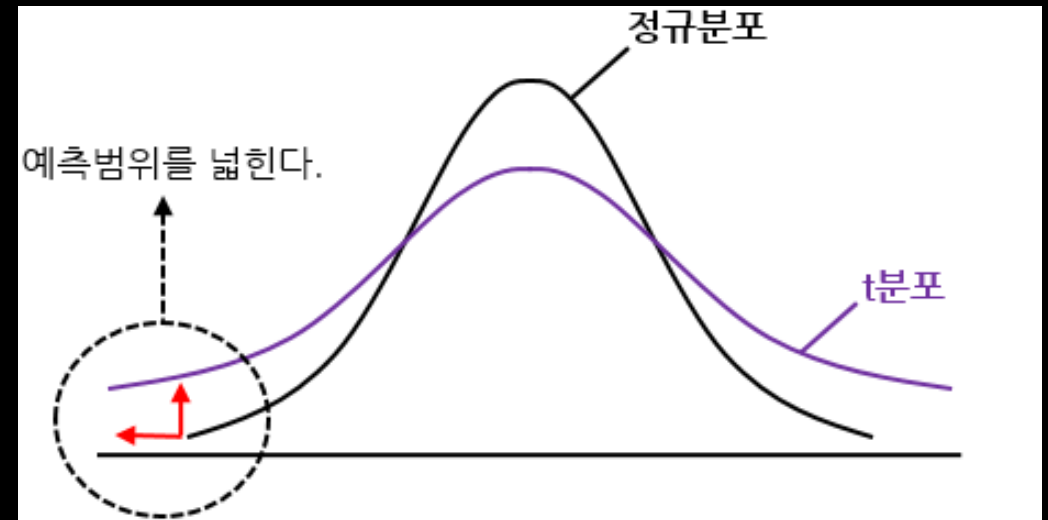
```
xtrans<-ecdf(x)(x)  
ytrans<-ecdf(y)(y)  
plot(xtrans,ytrans,xlab="CDF  
transfer",ylab="CDF  
transportation",main="transfer and  
transfortation CDFs")  
par(mfrow=c(1,1))
```



Transfer-Credit / t분포 적용

```
library(QRM)
tfit1 <- fit.st(x)
tpars1<-tfit1[[2]]
nu1<-tpars1[1]
mu1<-tpars1[2]
sigma1<-tpars1[3]

tfit2 <- fit.st(y)
tpars2 <- tfit2[[2]]
nu2<-tpars2[1]
mu2<-tpars2[2]
sigma2<-tpars2[3]
```

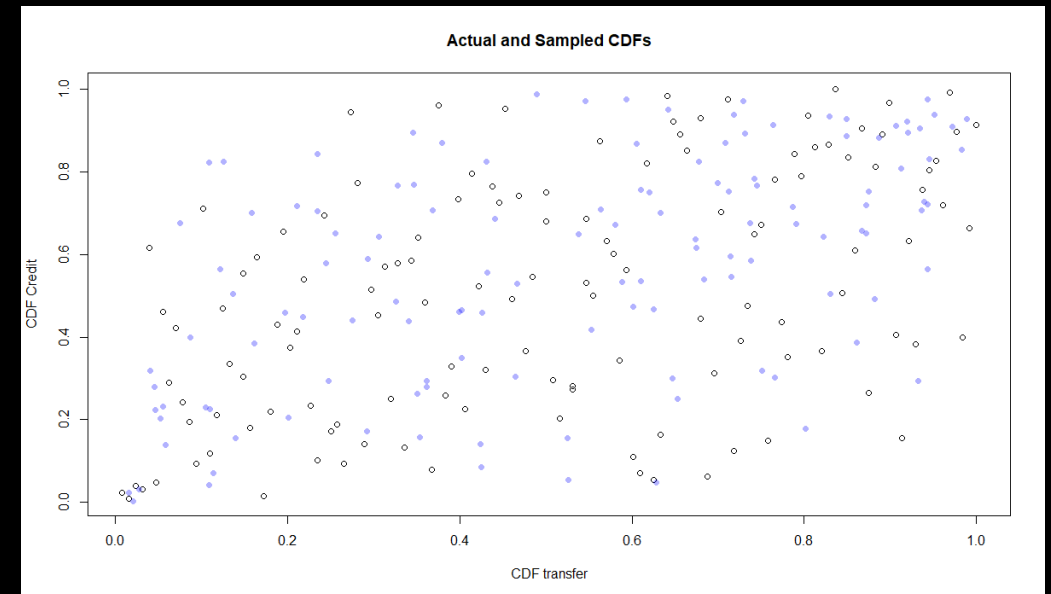


Transfer-Credit / 적합한 copula 함수 추정 및 적합성 판단

```
library(VineCopula)
FittedCopula<-
BiCopSelect(xtrans,ytrans,familyset=NA)
summary(FittedCopula) # Survival Gumbel

plot(xtrans,ytrans,xlab="CDF Credit",ylab="CDF
transfer",main="Actual and Sampled CDFs")

CopulaSim<-BiCopSim(length(x),
family=FittedCopula, par=FittedCopula)
points(CopulaSim, col = rgb(0,0,1,alpha=0.3),
pch=16)
```



Transfer-Credit / 99% 신뢰 수준 하 이용자 감소 리스크 측정

```
library(copula)
library(VC2copula)
w1<-0.5 w2<-0.5 #서비스 가중치 부여
n<-1000000
Percentiles<-c()
for(i in 1:20)
{sim<-
rCopula(n,copulaFromFamilyIndex(family=14,par=1
.46))
sim2<-
cbind(qt(sim[,1],df=nu1)*sigma1+mu1,qt(sim[,2],
df=nu2)*sigma2+mu2)
Percentiles<-c(Percentiles,sort(sim2 %*%
c(w1,w2))[0.01*n])}
VaR<-mean(Percentiles)
VaR
```

0.4465387

Biz Insight

고맙습니다.

