```
##線形回帰分析の拡張:一般化線形回帰
#farawayパッケージが入っていない場合はinstall.packages("faraway")で
#パッケージをインストールしておく.
data(gala,package="faraway")
# ガラパゴス諸島の30の島と亀の種類との関連
#7変数30サンプル
#Species:その島の亀の種類の数(従属変数)
#Endemics: 亀固有種の数 (説明変数)
#Area:島の面積(km2) (説明変数)
#Elevation:島の標高(m) (説明変数)
#Nearest:最近隣の島との距離(km)(説明変数)
#Scruz: Santa Cruz島との距離(km2) (説明変数)
#Adjacent: 近隣の島のエリア (km2) (説明変数)
str(gala);
plot(sort((gala$Species))) #Speciesは正の整数
gala.pm1<-glm(Species~ ., data = gala, family = poisson)</pre>
summary(gala.pm1)
gala.pm2<-step(gala.pm1) #AICで変数選択
summary(gala.pm2) #4変数+切片のモデルが選ばれる
confint(gala.pm2) #信頼区間の取り出し
c(AIC(gala.pm1),AIC(gala.pm2)) #AICを比較
#予測値の取り出し
gala.pred <- predict(gala.pm2, type="response")</pre>
#予測値と観測値のプロット
sortedgala <- sort(gala.pred,index.return = TRUE)</pre>
par(new=F)
plot(gala$Species[sortedgala$ix],pch=1,xlab="index",ylab="Species")
par(new=T)
lines(gala.pred[sortedgala$ix],col="red",pch=2,lty=1)
Iegend(5,400,c("予測値","観測値"),pch=c(-1,1),coI=2:1,Ity=c(1,0),cex=1.5)
#負の二項分布
#電子回路のはんだ不良箇所個数に関するデータ
data(solder,package="faraway")
str(solder) # 900x6
plot(sort(log(solder$skips))) #skipsは整数値
mod.nb<-glm(skips~.,family= negative.binomial(4),data=solder) #パラメータk=4の負の二項分布
summary(mod.nb)
#今回はdevianceがそこまで大きくなく.よい当てはまりといえる.
#負の二項分布のパラメータをAICで決めてみる.
#パラメータ数が同じなので最尤推定に相当
for(i in c(1:10)){
 cat(c(i,":"))
 cat(AIC(gIm(skips~.,family= negative.binomial(i),data=solder)))
 cat("\n")
```

```
}
# i=4でAIC最小
#ポアソン分布を用いたときのAICとの比較.
AIC(glm(skips~.,family= poisson(),data=solder))
#i=4における負の二項分布よりAICが大きい.
#よって負の二項分布(4)を採択
#予測値を取り出す
mod.pred <- predict(mod.nb,type="response");</pre>
#予測値と観測値をプロット
sortedskips <- sort(mod.pred,index.return = TRUE)</pre>
par(new=F)
plot(solder$skips[sortedskips$ix],ylab="skips")
par(new=T)
lines(mod.pred[sortedskips$ix],col="red",lwd="2")
legend(5,43,c("予測値","観測値"),pch=c(-1,1),col=2:1,lty=c(1,0),cex=1.5)
#logスケールでプロット
par(new=F)
plot(log(mod.pred[sortedskips$ix]),col="red",ylab="skips")
par(new=T)
points(log(solder$skips[sortedskips$ix]))
```