Risk premium using GLMs, QR, QRII, PQR and ER

Table of Contents

## Load packages

library(data.table)  
 library(gamlss)  
 library(quantreg)  
 library(qrcm)  
 library(knitr) # kable  
 library(cplm) # cpqlm  
 library(tweedie)  
 library(expectreg)  
  
# define the logistic link function with exposures  
 logitexp <- function(exposure = 1){  
 linkfun <- function(mu) {   
 eta <- if (any(exposure-mu <= 0)) log((mu)/abs(mu-exposure)) else log((mu)/(exposure-mu))  
 eta   
 }  
   
 linkinv <- function(eta) {  
 thresh <- -log(.Machine$double.eps)  
 eta <- pmin(thresh, pmax(eta, -thresh))  
 exposure\*(exp(eta)/(1 + exp(eta)))  
 }  
   
 mu.eta <- function(eta) {  
 thresh <- -log(.Machine$double.eps)  
 res <- rep(.Machine$double.eps, length(eta))  
 res[abs(eta) < thresh] <- ((exposure\*exp(eta))/(1 + exp(eta))^2)[abs(eta) < thresh]  
 res  
 }  
   
 valideta <- function(eta) TRUE   
 link <- paste("logexp(", exposure, ")", sep="")   
 structure(list(linkfun = linkfun, linkinv = linkinv, mu.eta = mu.eta,   
 valideta = valideta, name = link), class = "link-glm")   
 }

## Section 4: Simulation study

### Case: MSE and sample variance of the estimated risk premium with respect to the values of the risk loading parameter

set.seed(111)  
Nsim <- 2000  
mse.mat <- matrix(0, nrow = Nsim, ncol = 4) # mse results  
colnames(rmse.mat) <- c('QR', 'QRCF', "QRII" ,'EQR') # the competing models  
alpha.vector <- seq(from = 0, to = 0.15, by = 0.02) # risk loading parameters  
alpha.vector <- alpha.vector[-1]  
pars <- c(120, 1.65) # power and dispersion parameter in tweedie distribution  
  
mse <- svar.mse <- matrix(NA, nrow = length(alpha.vector), 4)  
  
for (j in 1:length(alpha.vector)){  
 alpha <- alpha.vector[j]  
 for(k in 1:Nsim){  
 nsample <- 5000  
 x1sim <- rbinom(n = nsample, size = 1, prob = 0.6) # covariates  
 x2sim <- rbinom(n = nsample, size = 1, prob = 0.8)  
 x3sim <- rbinom(n = nsample, size = 1, prob = 0.5)  
 beta <- c(4, 0.5, 0.5, 0.5) # coffieients  
 X <- model.matrix( ~ x1sim + x2sim + x3sim) # Regression coefficients  
 mu <- as.vector(exp(X%\*%beta)) # mean parameter in tweedie distribution  
   
   
 phi <- pars[1] # dispersion parameter  
 p <- pars[2] # power parameter  
 # true value of C, Pure premium, and Risk prmeium  
 C <- sum((1 + alpha)\*mu) # risk loading parameter is 10% using expected value premium principle  
 E.L <- sum(mu) # true pure premium  
 riskP <- mu\*(1 + alpha) # true risk premium for dt  
 ysim <- 0  
 for(i in 1:nsample){  
 ysim[i] <- rtweedie(n = 1, mu = mu[i], phi = phi, power = p)  
 }  
   
 dt <- data.table(x1 = x1sim,  
 x2 = x2sim,  
 x3 = x3sim,  
 y = ysim,  
 clm = ifelse(ysim == 0, 0, 1))  
 dt0 <- dt[y > 0]  
 dt$count <- 1  
 dt.sum <- dt[, .( y = sum(y), Npolicies = sum(count)),  
 by = c('x1', 'x2', 'x3')  
 ]  
 dt.sum <- dt.sum[order(x1, x2, x3)] # tariff classes for prediction  
 # logistic regression  
 m\_clm <- glm(clm ~ x1 + x2 + x3, family = binomial(link = "logit"), data = dt)  
 X1 <- model.matrix( ~ x1 + x2 + x3, data = dt.sum)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X1%\*%beta\_pred  
 pi.pred <- 1 - (exp(eta\_pred)/(1+exp(eta\_pred)))  
 dt.sum$pi.pred <- pi.pred  
 dt.sum$mu <- exp(X1%\*%beta)  
 # true risk premium   
 dt.sum$riskP <- (1 + alpha)\*dt.sum$mu  
  
 ## =======================================================================  
 # Two-part GLMs: Logistic + gamma  
 ## =======================================================================  
 # 0-1 claim or not part  
 m\_clm <- glm(clm ~ x1 + x2 + x3, family = binomial(link = "logit"), data = dt)  
 X\_pred <- model.matrix( ~ x1 + x2 + x3, data = dt)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X\_pred%\*%beta\_pred  
 pi\_pred <- 1 - (exp(eta\_pred)/(1+exp(eta\_pred)))#\*dt$exposure  
   
 # claim > 0 claim amount part  
 m.claim <- glm(y ~ x1 + x2 + x3, family = Gamma(link='log'), data = dt0)  
 X\_pred <- model.matrix( ~ x1 + x2 + x3, data = dt)  
 beta\_pred <- coef(m.claim)  
 y\_pred <- exp(X\_pred%\*%beta\_pred)  
   
   
   
 #-----------------------------  
 # prediction for dt.sum  
 #-----------------------------  
 X1 <- model.matrix( ~ x1 + x2 + x3, data = dt.sum)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X1%\*%beta\_pred  
 pi.pred <- 1 - (exp(eta\_pred)/(1 + exp(eta\_pred))) # non-claim probability (exposure=1)  
 beta\_pred <- coef(m.claim)  
 mu <- exp(X1%\*%beta\_pred)  
 E.ga <- (1 - pi.pred)\*mu  
   
 dt.sum$Pure <- E.ga # pure premium  
   
   
 ## =======================================================================  
 # Two-part quantile regressions  
 # QR & PQR(QRCF) & ER  
 ## =======================================================================  
   
 theta <- 0.95 # quantile level  
 dt.sum$theta\_star <- (theta - dt.sum$pi.pred)/(1 - dt.sum$pi.pred) # non-zero quantile level  
 #-----------------------------  
 # QR model  
 #-----------------------------  
 dt$QR.95 <- 0  
 for(i in 1:nrow(dt.sum)){  
 m.s <- rq(log(y) ~ x1 + x2 + x3, data = dt0, tau = dt.sum$theta\_star[i]) # QR model  
 TraQuan <- exp(predict(m.s, newdata = dt.sum))#\*(1-dt.sum$pi.pred) # prediction  
 dt.sum$QR.95[i] <- TraQuan[i]  
 }  
 qr.alpha <<- (C - E.L)/(sum(dt.sum$QR.95\*dt.sum$Npolicies) - E.L)   
   
 #-----------------------------  
 # QR  
 # Risk class prediction (dt.sum)   
 #-----------------------------  
 dt.sum$QR <- dt.sum$Pure + qr.alpha\*(dt.sum$QR.95 - dt.sum$Pure)   
 dt.sum$loading.QR <- dt.sum$QR - dt.sum$Pure   
   
   
 #-----------------------------  
 # QRII model  
 # Risk class prediction (dt.sum)   
 #-----------------------------  
 theta\_solve <- function(theta){  
 m.s <- rq(log(y) ~ x1 + x2 + x3, data = dt0, tau = theta)   
 flag <- exp(predict(m.s, newdata = dt.sum))\*(1-dt.sum$pi.pred)  
 bias <- C - sum(flag\*dt.sum$Npolicies)  
 return(bias)  
 }  
 upper\_theta <- 0.99999  
 lower\_theta <- 0.00000001  
 library(rootSolve)  
 # solve the quantile level (risk loading parameter in QRII)  
 m\_solve <- uniroot(f = theta\_solve,  
 lower = lower\_theta,  
 upper = upper\_theta,  
 maxiter = 50000,  
 tol = .Machine$double.eps^0.525)   
 theta\_Tra <- m\_solve$root  
 m.s <- rq(log(y) ~ x1 + x2 + x3, data = dt0, tau = theta\_Tra)   
 flag <- exp(predict(m.s, newdata = dt.sum))\*(1-dt.sum$pi.pred)  
 dt.sum$QRII <- flag  
   
 #-----------------------------  
 # PQR(QRCF) model  
 # solve risk loading parameter   
 #-----------------------------  
 m3.s <- iqr(log(y) ~ x1 + x2 + x3,  
 data = dt0,  
 formula.p = ~ slp(p,3))  
 dt.sum$QRCF.95 <- 0  
 for(i in 1:nrow(dt.sum)){  
 flag <- predict(m3.s, type = "QF", p = dt.sum$theta\_star[i], newdata = dt.sum)  
 dt.sum$QRCF.95[i] <- exp(flag$fit[i,1])  
 }  
 qrcf.alpha <<- (C - E.L)/(sum(dt.sum$QRCF.95\*dt.sum$Npolicies) - E.L) # risk loading parameter   
   
 #-----------------------------  
 # QRCF  
 # dt.sum  
 #-----------------------------  
 dt.sum$QRCF <- dt.sum$Pure + qrcf.alpha\*(dt.sum$QRCF.95 - dt.sum$Pure) # risk premium   
 dt.sum$loading.QRCF <- dt.sum$QRCF - dt.sum$Pure # risk loading   
   
   
   
 #-----------------------------  
 # expectiel regression (EQR)  
 # solve alpha  
 #-----------------------------  
   
 dt.sum$EQR.95 <- 0  
   
 m3.e <- expectreg.ls((y) ~ x1 + x2 + x3, data = dt, expectiles = 0.95)  
 flag <- predict(m3.e, newdata = dt.sum)  
 dt.sum$EQR.95 <- (flag$fit)  
 eqr.alpha <- (C - E.L)/(sum(dt.sum$EQR.95\*dt.sum$Npolicies) - E.L)   
   
   
   
 #-----------------------------  
 # EQR  
 # dt.sum  
 #-----------------------------  
 dt.sum$EQR <- dt.sum$Pure + eqr.alpha\*(dt.sum$EQR.95 - dt.sum$Pure)   
 dt.sum$loading.EQR <- dt.sum$EQR - dt.sum$Pure   
   
 # #-----------------------------  
 # # MSE  
 # #-----------------------------  
   
 mse.mat[k,] <- with(dt.sum, {  
 c(sum(((QR - riskP))^2),  
 sum(((QRCF - riskP))^2),  
 sum(((QRII - riskP))^2),  
 sum(((EQR - riskP))^2))/nrow(dt.sum)  
 })  
  
 }  
 mean.mse <- apply(mse.mat, 2, mean)  
 # mse for all models  
 mse[j,] <- apply(mse.mat, 2, mean)  
 # sample variance for all models  
 svar.mse[j,] <- c(sum((mse.mat[,1] - mean.mse[1])^2)/(Nsim - 1)/Nsim,  
 sum((mse.mat[,2] - mean.mse[2])^2)/(Nsim - 1)/Nsim,  
 sum((mse.mat[,3] - mean.mse[3])^2)/(Nsim - 1)/Nsim,  
 sum((mse.mat[,4] - mean.mse[4])^2)/(Nsim - 1)/Nsim)  
  
   
}

load("D:/对外经济贸易大学 - 科研/0. Risk Loadings of Classification Ratemaking（in progress)/2021-final-code/mse\_overall.RData")  
  
row.names(mse) <- alpha.vector  
row.names(svar.mse) <- alpha.vector  
colnames(mse) <- c('QR', 'QRCF', "QRII" ,'EQR')  
colnames(svar.mse) <- c('QR', 'QRCF', "QRII" ,'EQR')  
  
kable(mse, digits = 2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | QR | QRCF | QRII | EQR |
| 0.02 | 361.51 | 360.91 | 435.41 | 358.29 |
| 0.04 | 373.78 | 372.54 | 429.03 | 367.57 |
| 0.06 | 379.10 | 377.62 | 447.44 | 370.32 |
| 0.08 | 392.06 | 388.98 | 476.18 | 377.90 |
| 0.1 | 403.41 | 398.64 | 473.18 | 384.42 |
| 0.12 | 427.02 | 421.43 | 498.18 | 405.02 |
| 0.14 | 444.63 | 436.40 | 507.09 | 412.92 |

kable(svar.mse, digits = 2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | QR | QRCF | QRII | EQR |
| 0.02 | 36.16 | 36.03 | 72.75 | 35.49 |
| 0.04 | 43.27 | 43.07 | 69.21 | 41.97 |
| 0.06 | 42.74 | 42.63 | 81.84 | 41.14 |
| 0.08 | 45.58 | 45.01 | 89.61 | 41.80 |
| 0.1 | 47.01 | 46.32 | 87.49 | 43.38 |
| 0.12 | 54.48 | 53.31 | 89.84 | 49.11 |
| 0.14 | 63.44 | 61.70 | 104.13 | 54.30 |

### Case: MSE and sample variance of the estimated risk premium for 8 tariff classes in the case of .

load("D:/对外经济贸易大学 - 科研/0. Risk Loadings of Classification Ratemaking（in progress)/2021-final-code/mse\_riskclass.RData")  
mse.caseI <- rbind(mse1[1,], mse2[1,], mse3[1,], mse4[1,], mse5[1,], mse6[1,], mse7[1,], mse8[1,])  
mse.caseII <- rbind(mse1[2,], mse2[2,], mse3[2,], mse4[2,], mse5[2,], mse6[2,], mse7[2,], mse8[2,])  
mse.caseIII <- rbind(mse1[3,], mse2[3,], mse3[3,], mse4[3,], mse5[3,], mse6[3,], mse7[3,], mse8[3,])  
colnames(mse.caseI) <- c('QR', 'QRCF', "QRII" ,'EQR')  
colnames(mse.caseII) <- c('QR', 'QRCF', "QRII" ,'EQR')  
colnames(mse.caseIII) <- c('QR', 'QRCF', "QRII" ,'EQR')  
row.names(mse.caseI) <- riskclass  
row.names(mse.caseII) <- riskclass  
row.names(mse.caseIII) <- riskclass  
mse.all <- rbind(mse.caseI, mse.caseII, mse.caseIII)  
# mse  
kable(mse.all, digits = 2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | QR | QRCF | QRII | EQR |
| 000 | 122.29 | 122.90 | 197.79 | 129.29 |
| 001 | 312.93 | 312.61 | 480.75 | 309.81 |
| 010 | 151.48 | 150.96 | 230.59 | 150.08 |
| 011 | 401.49 | 398.00 | 486.41 | 388.02 |
| 100 | 298.48 | 297.42 | 447.28 | 293.37 |
| 101 | 751.21 | 745.01 | 1022.79 | 715.89 |
| 110 | 294.80 | 292.88 | 331.60 | 287.23 |
| 111 | 713.38 | 712.08 | 451.70 | 704.46 |
| 000 | 153.25 | 153.22 | 217.13 | 169.93 |
| 001 | 374.04 | 371.76 | 521.01 | 367.09 |
| 010 | 178.00 | 176.06 | 242.34 | 173.14 |
| 011 | 426.94 | 420.54 | 483.85 | 400.45 |
| 100 | 329.62 | 327.83 | 449.54 | 323.33 |
| 101 | 811.47 | 802.39 | 1035.35 | 746.70 |
| 110 | 320.93 | 316.72 | 355.94 | 301.91 |
| 111 | 729.27 | 721.87 | 494.19 | 703.28 |
| 000 | 164.02 | 165.21 | 241.37 | 193.22 |
| 001 | 395.15 | 393.90 | 562.78 | 386.94 |
| 010 | 202.52 | 197.83 | 274.43 | 194.58 |
| 011 | 458.26 | 448.35 | 532.89 | 415.47 |
| 100 | 352.95 | 349.32 | 490.51 | 337.49 |
| 101 | 884.55 | 868.81 | 1163.47 | 783.64 |
| 110 | 327.06 | 320.44 | 358.63 | 301.07 |
| 111 | 766.33 | 759.63 | 538.33 | 734.76 |

## Section 5: Case study: Australian automobile insurance

### read data set from the local disk

dt <- fread('C:/Users/Administrator/Desktop/expectile-regression-risk-loading/car.csv')  
 dt$claimcst00 <- dt$claimcst0/dt$exposure  
 dt$count <- 1  
 dt$veh\_age <- factor(dt$veh\_age, levels = c(2, 1, 3, 4))  
 dt$agecat <- factor(dt$agecat, levels = c(5, 1, 2, 3, 4, 6))  
 dt0 <- subset(dt, claimcst0 > 0)  
 # data set for 24 risk classes  
 dt.sum <- dt[, .(Ncount = sum(count),   
 exposure = sum(exposure),   
 Npolicies = sum(count)),  
 by = c('veh\_age', 'agecat')]  
   
 # modelling claim probability  
 m\_clm <- glm(clm ~ veh\_age + agecat, family=binomial(link=logitexp(dt$exposure)), data=dt)  
 X1 <- model.matrix( ~ veh\_age + agecat, data=dt.sum)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X1%\*%beta\_pred  
 pi.pred <- 1 - (exp(eta\_pred)/(1+exp(eta\_pred)))   
 dt.sum$pi.pred <- pi.pred  
 dt.sum <- dt.sum[order(pi.pred)] # sort by the claim probability  
 kable(dt.sum, digits = 3)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| veh\_age | agecat | Ncount | exposure | Npolicies | pi.pred |
| 2 | 1 | 1504 | 697.437 | 1504 | 0.798 |
| 1 | 1 | 1283 | 556.038 | 1283 | 0.803 |
| 3 | 1 | 1643 | 748.348 | 1643 | 0.818 |
| 2 | 2 | 3167 | 1486.946 | 3167 | 0.828 |
| 4 | 1 | 1312 | 610.450 | 1312 | 0.831 |
| 1 | 2 | 2160 | 937.520 | 2160 | 0.833 |
| 2 | 3 | 3741 | 1770.231 | 3741 | 0.837 |
| 1 | 3 | 2706 | 1194.021 | 2706 | 0.841 |
| 2 | 4 | 3919 | 1880.767 | 3919 | 0.843 |
| 3 | 2 | 3956 | 1820.019 | 3956 | 0.846 |
| 1 | 4 | 2935 | 1258.995 | 2935 | 0.847 |
| 3 | 3 | 4826 | 2294.689 | 4826 | 0.853 |
| 4 | 2 | 3592 | 1647.387 | 3592 | 0.857 |
| 3 | 4 | 4760 | 2314.018 | 4760 | 0.859 |
| 4 | 3 | 4494 | 2150.516 | 4494 | 0.865 |
| 4 | 4 | 4575 | 2162.762 | 4575 | 0.870 |
| 2 | 5 | 2635 | 1296.041 | 2635 | 0.871 |
| 2 | 6 | 1621 | 792.255 | 1621 | 0.871 |
| 1 | 5 | 2042 | 913.057 | 2042 | 0.874 |
| 1 | 6 | 1131 | 479.321 | 1131 | 0.875 |
| 3 | 5 | 3088 | 1508.583 | 3088 | 0.884 |
| 3 | 6 | 1791 | 856.454 | 1791 | 0.885 |
| 4 | 5 | 2971 | 1453.328 | 2971 | 0.894 |
| 4 | 6 | 2004 | 971.636 | 2004 | 0.894 |

### total risk premium for the whole insurance portfolio

C <- 22206147 # total risk premium that the insurer should charge

### Two-part GLMs: Logistic + gamma (EVPP + SDPP)

# 0-1 claim or not part  
 m\_clm <- glm(clm ~ veh\_age + agecat, family=binomial(link=logitexp(exposure=dt$exposure)), data=dt)  
 X\_pred <- model.matrix( ~ veh\_age + agecat, data = dt)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X\_pred%\*%beta\_pred  
 pi\_pred <- 1 - (exp(eta\_pred)/(1+exp(eta\_pred)))#\*dt$exposure  
  
 # claim > 0 claim amount part  
 m.claim <- glm(claimcst0 ~ veh\_age + agecat, family = Gamma(link='log'), data = dt0)  
 X\_pred <- model.matrix( ~ veh\_age + agecat, data = dt)  
 beta\_pred <- coef(m.claim)  
 y\_pred <- exp(X\_pred%\*%beta\_pred)  
   
 #--------------------------------------------------------------------------  
 # solve the risk loading parameter in EVPP and SDPP: alpha.ga.E and alpha.ga.sd  
 #--------------------------------------------------------------------------  
 Dis\_par <- 3.103186 # Dispersion parameter  
 E.ga <- (1 - pi\_pred)\*y\_pred  
 Var.ga <- (1 - pi\_pred)\*y\_pred^2\*(pi\_pred + Dis\_par)  
 sd.ga <- sqrt(Var.ga)  
 alpha.ga.E <- C/sum(E.ga) - 1 # risk roading parameter in EVPP  
 alpha.ga.sd <- (C - sum(E.ga))/sum(sd.ga) # risk roading parameter in SDPP  
 c(alpha.ga.E\*100, alpha.ga.sd\*100)

## [1] 11.966303 2.309759

#----------------------------------------------------------  
 # estimation of risk premium for 24 risk classes (dt.sum)  
 #----------------------------------------------------------  
 X1 <- model.matrix( ~ veh\_age + agecat, data = dt.sum)  
 beta\_pred <- coef(m\_clm)  
 eta\_pred <- X1%\*%beta\_pred  
 pi.pred <- 1 - (exp(eta\_pred)/(1 + exp(eta\_pred)))  
   
 beta\_pred <- coef(m.claim)  
 mu <- exp(X1%\*%beta\_pred)  
 E.ga <- (1 - pi.pred)\*mu  
 var.ga <- (1 - pi.pred)\*mu^2\*(pi.pred + Dis\_par)  
 sd.ga <- sqrt(var.ga)  
  
 dt.sum$Pure <- E.ga # pure premium  
 dt.sum$SDPP <- E.ga + alpha.ga.sd\*sd.ga # risk premium in EVPP  
 dt.sum$EVPP <- E.ga + alpha.ga.E\*E.ga # risk premium in SDPP  
  
 dt.sum$loading.E <- dt.sum$EVPP - dt.sum$Pure # risk loading in EVPP  
 dt.sum$loading.sd <- dt.sum$SDPP - dt.sum$Pure # risk loading in EVPP

### QRII

theta\_Tra <- 0.7908  
 m.QRII <- rq(log(claimcst0) ~ veh\_age + agecat, data = dt0, tau = theta\_Tra)   
 dt.sum$QRII <- exp(predict(m.QRII, newdata = dt.sum))\*(1-dt.sum$pi.pred)

### QR

# solve the risk loading parameter  
 theta <- 0.95 # quantile level for aggregate claim amount  
 dt.sum$theta\_star <- (theta - dt.sum$pi.pred)/(1 - dt.sum$pi.pred) # quantile level for non-zero aggregate claim amount  
 E.L <- sum(dt.sum$Pure\*dt.sum$Npolicies) # total pure premium  
 dt.sum$QR.95 <- rep(NA,24)  
 for(i in 1:24){  
 m.s <- rq(log(claimcst0) ~ veh\_age + agecat, data = dt0, tau = dt.sum$theta\_star[i]) # QR model  
 TraQuan <- exp(predict(m.s, newdata = dt.sum))  
 dt.sum$QR.95[i] <- TraQuan[i]  
 }  
 qr.alpha <- (C - E.L)/(sum(dt.sum$QR.95\*dt.sum$Npolicies) - E.L) # risk loading parameter  
 qr.alpha

## [1] 0.03001939

# Risk class prediction (dt.sum)   
 dt.sum$QR <- dt.sum$Pure + qr.alpha\*(dt.sum$QR.95 - dt.sum$Pure)   
 dt.sum$loading.QR <- dt.sum$QR - dt.sum$Pure

### PQR

# solve the risk loading parameter in QPP  
 m3.s <- iqr(log(claimcst0) ~ veh\_age + agecat,  
 data = dt0,  
 formula.p = ~ slp(p,3))  
   
 dt.sum$QRCF.95 <- rep(NA,24)  
 for(i in 1:24){  
 flag <- predict(m3.s, type = "QF", p = dt.sum$theta\_star[i], newdata = dt.sum)  
 dt.sum$QRCF.95[i] <- exp(flag$fit[i,1])  
 }  
 qrcf.alpha <- (C - E.L)/(sum(dt.sum$QRCF.95\*dt.sum$Npolicies) - E.L)   
 qrcf.alpha # the risk loading parameter

## [1] 0.03093399

# Risk class prediction (dt.sum)   
 dt.sum$PQR <- dt.sum$Pure + qrcf.alpha\*(dt.sum$QRCF.95 - dt.sum$Pure) # prediction in PQR for 24 tariff classes  
 dt.sum$loading.PQR <- dt.sum$PQR - dt.sum$Pure

### ER

# solve the risk loading parameter  
 library(expectreg) # expectreg.ls  
 library(dummies) # generate the dummies  
 ohe\_feats <- names(dt)[sapply(dt, is.factor)]  
 df\_all <- dt  
 df\_all\_sum <- dt.sum  
 for (f in ohe\_feats){  
 df\_all <- dummy.data.frame(df\_all, names = f, sep = "\_")  
 df\_all\_sum <- dummy.data.frame(df\_all\_sum, names = f, sep = "\_")  
 }  
  
 dt.sum$EQR.95 <- rep(NA,24)  
 m3.e <- expectreg.ls((claimcst0) ~ agecat\_1 + agecat\_2 + agecat\_3 + agecat\_4 + agecat\_6  
 + veh\_age\_1 + veh\_age\_3 + veh\_age\_4, data = df\_all,   
 expectiles = 0.95,  
 ci = T)  
 flag <- predict(m3.e, newdata = df\_all\_sum)  
 dt.sum$EQR.95 <- (flag$fit)  
 eqr.alpha <- (C - E.L)/(sum(dt.sum$EQR.95\*dt.sum$Npolicies) - E.L)   
 eqr.alpha # the risk loading parameter in EPP

## [1] 0.02845861

# Risk class prediction (dt.sum)   
 dt.sum$EQR <- dt.sum$Pure + eqr.alpha\*(dt.sum$EQR.95 - dt.sum$Pure) # prediction in EQR for individual polices  
 dt.sum$loading.EQR <- dt.sum$EQR - dt.sum$Pure # risk loading

### Risk premium comparsion

dt.out <- dt.sum[,c('veh\_age','agecat',"pi.pred",  
 'Pure', 'EVPP', 'SDPP',  
 'QR','QRII','PQR','EQR')]  
kable(dt.out, digits = 2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| veh\_age | agecat | pi.pred | Pure | EVPP | SDPP | QR | QRII | PQR | EQR |
| 2 | 1 | 0.80 | 522.88 | 585.45 | 575.97 | 603.63 | 728.58 | 602.91 | 578.98 |
| 1 | 1 | 0.80 | 484.58 | 542.56 | 534.43 | 546.13 | 585.84 | 550.14 | 535.93 |
| 3 | 1 | 0.82 | 484.98 | 543.01 | 536.93 | 557.52 | 771.13 | 562.59 | 538.73 |
| 2 | 2 | 0.83 | 355.42 | 397.95 | 394.69 | 396.57 | 415.44 | 396.54 | 396.64 |
| 4 | 1 | 0.83 | 491.20 | 549.97 | 546.00 | 564.34 | 784.62 | 570.13 | 546.14 |
| 1 | 2 | 0.83 | 329.07 | 368.45 | 365.94 | 361.44 | 333.73 | 362.93 | 365.21 |
| 2 | 3 | 0.84 | 302.31 | 338.49 | 336.62 | 339.95 | 381.75 | 339.79 | 338.52 |
| 1 | 3 | 0.84 | 279.83 | 313.31 | 312.03 | 311.27 | 306.58 | 310.90 | 310.84 |
| 2 | 4 | 0.84 | 296.12 | 331.55 | 330.36 | 327.68 | 346.93 | 329.20 | 332.39 |
| 3 | 2 | 0.85 | 328.44 | 367.75 | 366.80 | 369.35 | 438.08 | 367.72 | 367.01 |
| 1 | 4 | 0.85 | 274.05 | 306.84 | 306.18 | 300.14 | 278.58 | 301.50 | 305.11 |
| 3 | 3 | 0.85 | 279.08 | 312.47 | 312.56 | 315.36 | 402.14 | 314.74 | 312.52 |
| 4 | 2 | 0.86 | 331.81 | 371.52 | 372.21 | 373.98 | 444.62 | 371.18 | 371.65 |
| 3 | 4 | 0.86 | 273.17 | 305.86 | 306.58 | 303.50 | 365.21 | 304.57 | 306.67 |
| 4 | 3 | 0.86 | 281.74 | 315.45 | 316.99 | 317.41 | 407.84 | 317.31 | 316.47 |
| 4 | 4 | 0.87 | 275.65 | 308.63 | 310.80 | 306.73 | 370.22 | 306.85 | 310.44 |
| 2 | 5 | 0.87 | 215.94 | 241.78 | 243.59 | 239.92 | 273.48 | 238.77 | 244.89 |
| 2 | 6 | 0.87 | 234.28 | 262.31 | 264.32 | 261.66 | 257.69 | 259.32 | 264.52 |
| 1 | 5 | 0.87 | 199.67 | 223.56 | 225.60 | 218.81 | 219.41 | 219.07 | 223.25 |
| 1 | 6 | 0.87 | 216.63 | 242.55 | 244.79 | 238.56 | 206.73 | 238.00 | 241.53 |
| 3 | 5 | 0.88 | 198.53 | 222.28 | 225.45 | 220.03 | 286.91 | 219.78 | 224.55 |
| 3 | 6 | 0.88 | 215.38 | 241.15 | 244.62 | 240.96 | 270.33 | 239.05 | 242.73 |
| 4 | 5 | 0.89 | 199.86 | 223.77 | 228.15 | 220.55 | 290.16 | 220.19 | 227.21 |
| 4 | 6 | 0.89 | 216.81 | 242.76 | 247.55 | 242.19 | 273.39 | 239.68 | 245.49 |