```
Call:
glm(formula = dy_payer ~ dx_pay_count + dx_payer, family = "binomial",
    data = DTtrain)

Deviance Residuals:
    Min      1Q   Median       3Q      Max
-2.2198  -0.1228  -0.1228  -0.1228   3.1275

Coefficients:
              Estimate Std. Error  z value Pr(>|z|)
(Intercept)  -4.88318    0.04104 -118.972  < 2e-16 ***
dx_pay_count  0.45323    0.10275    4.411 1.03e-05 ***
dx_payerTRUE  4.53863    0.19722   23.013  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9543.4  on 80000  degrees of freedom
Residual deviance: 7599.5  on 79998  degrees of freedom
AIC: 7605.5

Number of Fisher Scoring iterations: 7
```
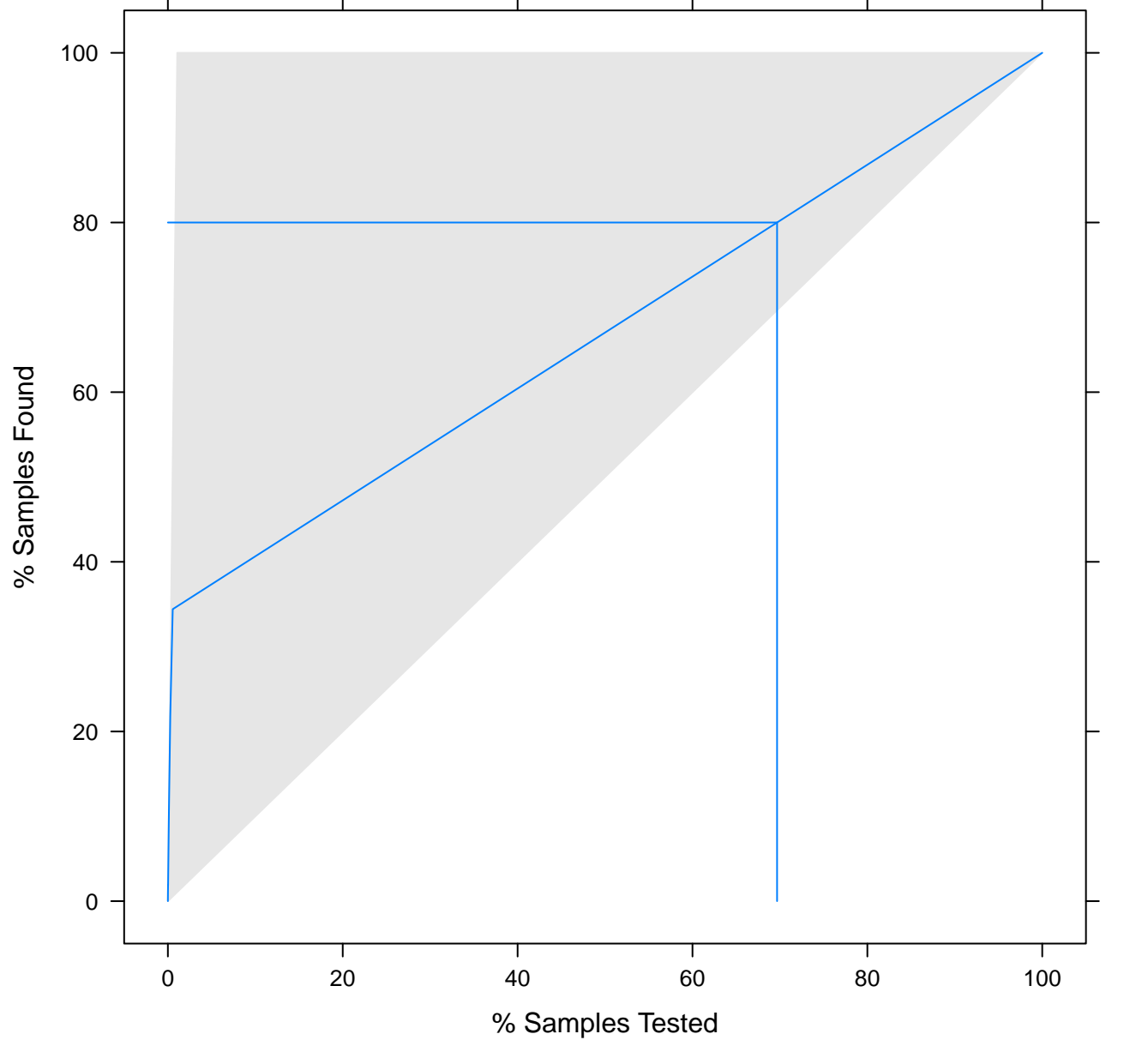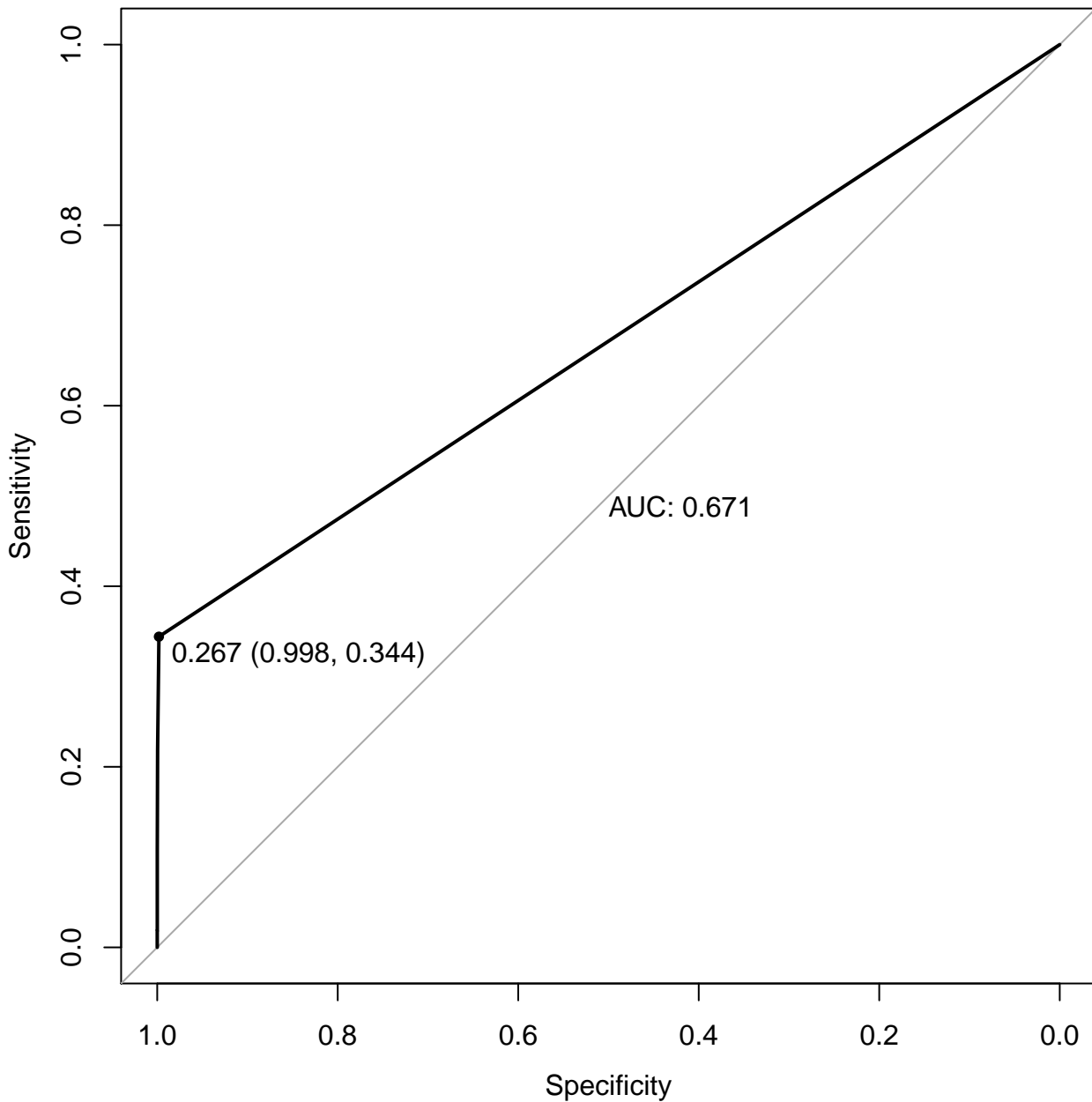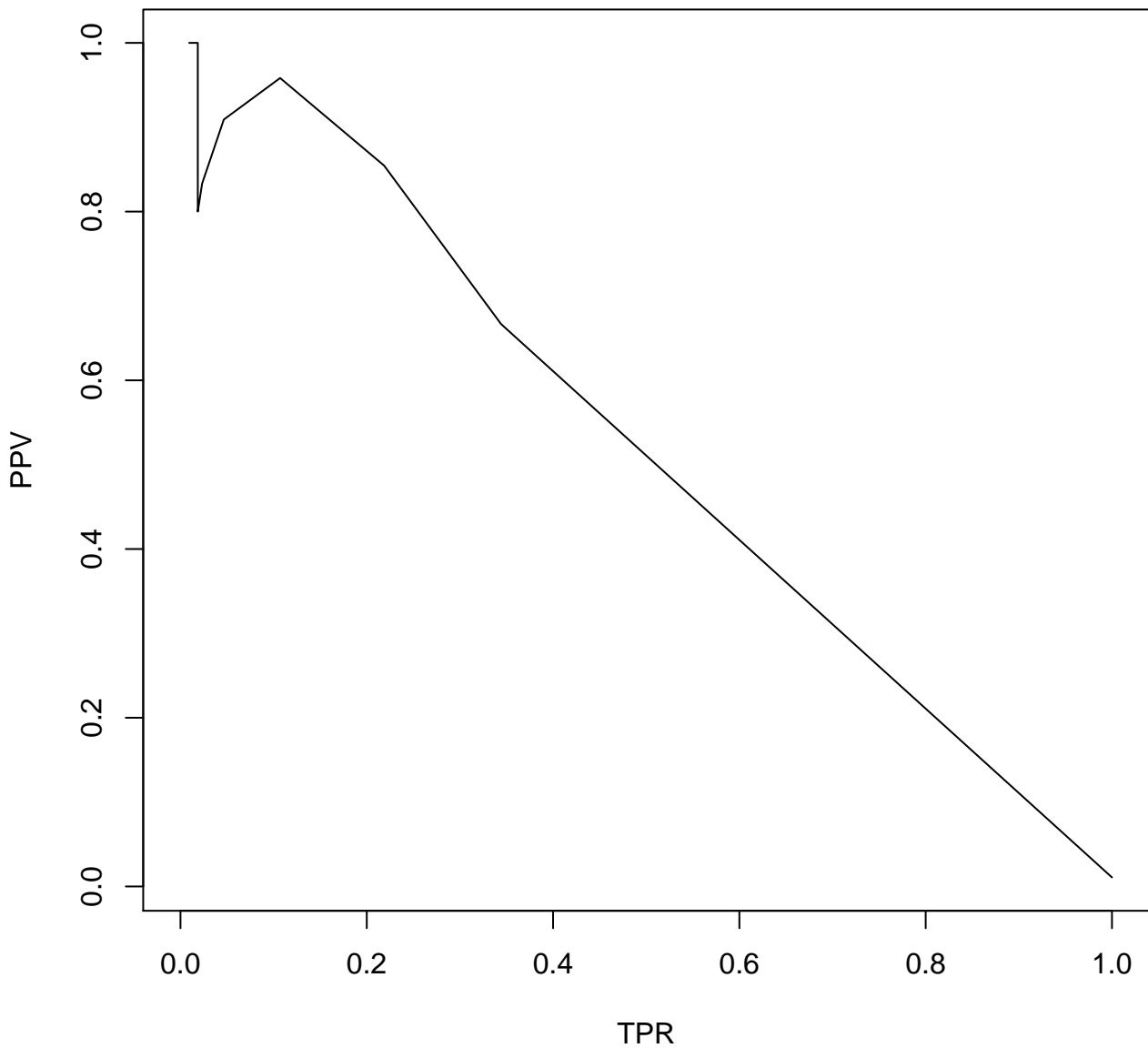
Lift chart

ROC curve
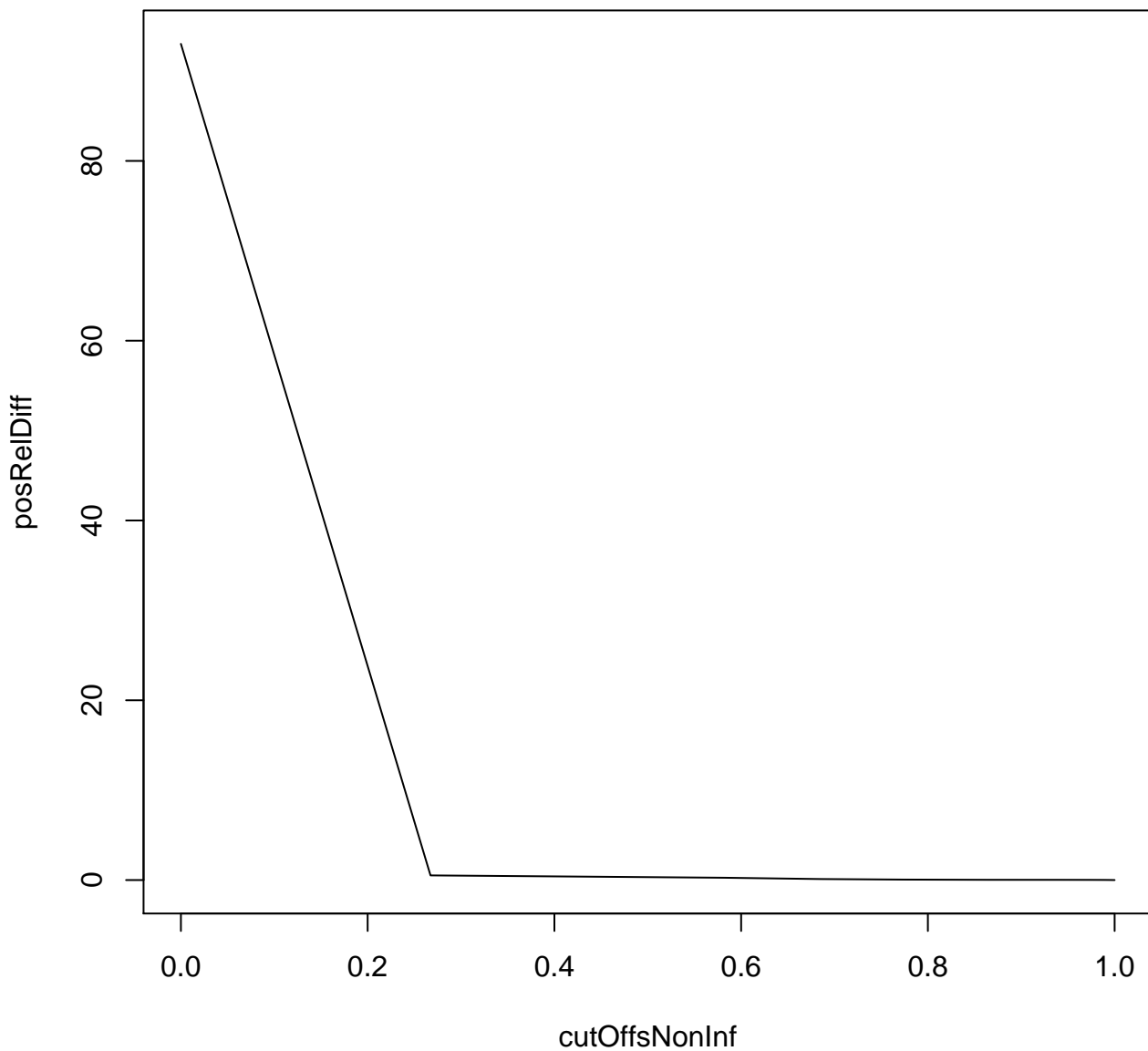
AUC: 0.671

0.267 (0.998, 0.344)

Specificity

Sensitivity

Sensitivity vs. Precision

predicted positives / actual positives

```
[1] "1st matrix – threshold based on the ROC curve"
[2] "0.26732928651714"
[3] "2nd matrix – threshold based on the predicted vs. actual count of positives"
[4] "0.26732928651714"
```

```
Confusion Matrix and Statistics

          Reference
Prediction FALSE  TRUE
     FALSE 19747   141
     TRUE     37    74

                 Accuracy : 0.9911
                   95% CI : (0.9897, 0.9924)
      No Information Rate : 0.9892
      P-Value [Acc > NIR] : 0.005133

                    Kappa : 0.45
 Mcnemar's Test P-Value : 1.162e-14

              Sensitivity : 0.34419
              Specificity : 0.99813
           Pos Pred Value : 0.66667
           Neg Pred Value : 0.99291
               Prevalence : 0.01075
           Detection Rate : 0.00370
     Detection Prevalence : 0.00555
        Balanced Accuracy : 0.67116

         'Positive' Class : TRUE
```

```
Confusion Matrix and Statistics

          Reference
Prediction FALSE  TRUE
     FALSE 19747   141
     TRUE     37    74

                Accuracy : 0.9911
                  95% CI : (0.9897, 0.9924)
     No Information Rate : 0.9892
     P-Value [Acc > NIR] : 0.005133

                   Kappa : 0.45
 Mcnemar's Test P-Value : 1.162e-14

             Sensitivity : 0.34419
             Specificity : 0.99813
          Pos Pred Value : 0.66667
          Neg Pred Value : 0.99291
              Prevalence : 0.01075
          Detection Rate : 0.00370
    Detection Prevalence : 0.00555
       Balanced Accuracy : 0.67116

        'Positive' Class : TRUE
```