

# Backward Elimination

Shan-Shan Song, *ss5422@columbia.edu*

12/10/2017

## Backward Elimination: “By-hand”

### Procedures

- start with all predictors in the model

```
# fit
bwd_fit_full = lm(log_length_of_stay ~ ., data = data_bwd)
arrange(tidy(bwd_fit_full), desc(p.value))
```

- remove the predictor with highest p-value
- refit the model and move on
- stop when all p-values  $\leq$  **alpha\_crit (p-to-remove) = 0.05**.

```
# step 1: no race
bwd_step1 = update(bwd_fit_full, . ~ . -race)
arrange(tidy(bwd_step1), desc(p.value))
step1_pvalue = tidy(bwd_step1) %>% select(term, p.value)
```

```
# step 2: no cindex
bwd_step2 = update(bwd_step1, . ~ . -cindex)
arrange(tidy(bwd_step2), desc(p.value))
step2_pvalue = tidy(bwd_step2) %>% select(term, p.value)
```

```
# step 3: no mews
bwd_step3 = update(bwd_step2, . ~ . -mews)
arrange(tidy(bwd_step3), desc(p.value))
step3_pvalue = tidy(bwd_step3) %>% select(term, p.value)
```

```
# step 4: no religion
bwd_step4 = update(bwd_step3, . ~ . -religion)
arrange(tidy(bwd_step4), desc(p.value))
step4_pvalue = tidy(bwd_step4) %>% select(term, p.value)
```

```
# step 5: no o2sat_cat
bwd_step5 = update(bwd_step4, . ~ . -o2sat_cat)
arrange(tidy(bwd_step5), desc(p.value))
step5_pvalue = tidy(bwd_step5) %>% select(term, p.value)
```

```
# step 6: no icu_flag
bwd_step6 = update(bwd_step5, . ~ . -icu_flag)
arrange(tidy(bwd_step6), desc(p.value))
step6_pvalue = tidy(bwd_step6) %>% select(term, p.value)
```

```
# step 7: no temperature_cat
bwd_step7 = update(bwd_step6, . ~ . -temperature_cat)
arrange(tidy(bwd_step7), desc(p.value))
step7_pvalue = tidy(bwd_step7) %>% select(term, p.value)
```

Table 1: P-values of Backward Elimination

term	pvalue_step1	pvalue_step2	pvalue_step3	pvalue_step4	pvalue_step5	pvalue_step6	pvalue_step7	pvalue_step8
(Intercept)	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
age	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
bmi	0.1100052	0.1137921	0.1059360	0.1497398	0.1708469	0.1771045	0.1593989	NA
cindexmoderate	0.0886317	NA	NA	NA	NA	NA	NA	NA
cindexnormal	0.9511007	NA	NA	NA	NA	NA	NA	NA
cindexsevere	0.0000243	NA	NA	NA	NA	NA	NA	NA
evisit	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
genderMale	0.0689479	0.0362905	0.0368188	0.0568089	0.0543859	0.0537196	0.0490663	0.0406418
heartrate_transformed	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
icu_flag1	0.3341491	0.3952503	0.4019109	0.4070222	0.4152766	NA	NA	NA
insurancetypeMedicare	0.0305833	0.0398744	0.0377183	0.0409622	0.0425393	0.0436914	0.0445049	0.0392080
insurancetypePrivate	0.0014031	0.0012652	0.0012498	0.0013847	0.0014109	0.0015382	0.0015715	0.0012439
is30dayreadmit1	0.0003639	0.0002384	0.0002305	0.0002325	0.0002529	0.0002585	0.0003848	0.0003512
maritalstatusNot Married	0.0296711	0.0306452	0.0305300	0.0200998	0.0195627	0.0204268	0.0359464	0.0307398
mean_arterial_pressure	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
mewsimmediate action	0.6482779	0.6835415	NA	NA	NA	NA	NA	NA
mewsincrease caution	0.7771244	0.9217115	NA	NA	NA	NA	NA	NA
mewsnormal	0.5974104	0.6445807	NA	NA	NA	NA	NA	NA
o2sat_catnormal	0.4674724	0.4399063	0.4348712	0.4295952	NA	NA	NA	NA
religionHindu	0.4538411	0.4345970	0.4461904	NA	NA	NA	NA	NA
religionIslam	0.0820389	0.1308588	0.1295581	NA	NA	NA	NA	NA
religionJewish	0.8191006	0.5825357	0.5920266	NA	NA	NA	NA	NA
religionNo Affiliation	0.0322253	0.0295137	0.0297650	NA	NA	NA	NA	NA
religionOther	0.1728072	0.2181224	0.2190891	NA	NA	NA	NA	NA
respirationrate_transformed	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
temperature_catlow	0.3147504	0.3446520	0.3438565	0.3732990	0.3626964	0.3631584	NA	NA
temperature_catnormal	0.0000232	0.0000314	0.0000292	0.0000315	0.0000301	0.0000342	NA	NA

```

# step 8: no bmi
bwd_step8 = update(bwd_step7, . ~ . -bmi)
arrange(tidy(bwd_step8), desc(p.value))
step8_pvalue = tidy(bwd_step8) %>% select(term, p.value)

# combine pvalues each step into a table
for (i in 1:8){
  if (i == 1){
    step_pvalue = step1_pvalue %>% rename(c("p.value" = "pvalue_step1"))
    next}
  loop_name = paste0("step",i,"_pvalue",sep = "")
  step_pvalue = merge(step_pvalue, eval(parse(text = loop_name)), all = TRUE, by = "term") %>%
    rename(c("p.value" = paste0("pvalue_step",i)))
}

```

## Results

```

# backward p-values
step_pvalue %>%
  knitr::kable(format = "latex",
               caption = "P-values of Backward Elimination",
               booktabs = TRUE) %>%
  kableExtra::kable_styling(latex_options = "scale_down")

# summary of model from backward elimination
bwd_step8 %>% summary()

```

```
##
## Call:
## lm(formula = log_length_of_stay ~ is30dayreadmit + evisit + age +
##     gender + maritalstatus + insurancetype + heartrate_transformed +
##     respirationrate_transformed + mean_arterial_pressure, data = data_bwd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.0662 -0.4702  0.0065  0.5045  3.2967
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.532e+00  1.971e-01  12.847 < 2e-16 ***
## is30dayreadmit  1.708e-01  4.773e-02   3.579 0.000351 ***
## evisit          6.943e-02  1.057e-02   6.569 6.07e-11 ***
## age            9.628e-03  9.928e-04   9.698 < 2e-16 ***
## genderMale      6.741e-02  3.291e-02   2.048 0.040642 *
## maritalstatusNot Married  7.075e-02  3.273e-02   2.162 0.030740 *
## insurancetypeMedicare -1.648e-01  7.986e-02  -2.063 0.039208 *
## insurancetypePrivate -2.438e-01  7.542e-02  -3.232 0.001244 **
## heartrate_transformed -2.173e+03  3.057e+02  -7.109 1.49e-12 ***
## respirationrate_transformed -2.048e+02  2.962e+01  -6.914 5.85e-12 ***
## mean_arterial_pressure -6.086e-03  9.802e-04  -6.210 6.13e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8251 on 2708 degrees of freedom
## (30 observations deleted due to missingness)
## Multiple R-squared:  0.1384, Adjusted R-squared:  0.1352
## F-statistic: 43.49 on 10 and 2708 DF, p-value: < 2.2e-16
```