Assignment 2

Michal Malyska

10/02/2020

```r
df1 <- read_rds(path = paste0(here(), "/data/ON_mortality.RDS")) %>%
    mutate(age = as.numeric(if_else(age == "110+", "110", age)))
```

Question 1

$$\lambda(t) = \alpha e^{\beta t}$$

a

$$S(t) = \exp(-\frac{\alpha}{\beta}(e^{\beta t} - 1))$$

Showing $S(t)$ (kind of reverse showing since I show that this $S(t)$ implies our hazard function but still valid.)

$$\begin{aligned}
\lambda(t) &= -\frac{d}{dt} log(S(t)) \\
&= -\frac{d}{dt}\left(-\frac{\alpha}{\beta}(e^{\beta t} - 1)\right) \\
&= \frac{d}{dt}\left(\frac{\alpha}{\beta}(e^{\beta t} - 1)\right) \\
&= \frac{d}{dt}\left(\frac{\alpha}{\beta}e^{\beta t}\right) \\
&= \alpha e^{\beta t}
\end{aligned}$$

$$f(t) = \alpha \exp(\beta t - \frac{\alpha}{\beta}(e^{\beta t} - 1))$$

Showing $f(t)$

$$\lambda(t) = \frac{f(t)}{S(t)} f(t) = \lambda(t)S(t)$$

$$\begin{aligned}
f(t) &= \lambda(t)S(t) \\
&= \alpha e^{\beta t} * \exp(-\frac{\alpha}{\beta}(e^{\beta t} - 1)) \\
&= \alpha \exp(\beta t - \frac{\alpha}{\beta}(e^{\beta t} - 1))
\end{aligned}$$

b

Modal Time of Death (mode of $f(t)$)

$$\frac{d}{dt}f(t) = f(t) * (\beta - \alpha e^{\beta t}) = 0$$
$$\implies (\beta - \alpha e^{\beta t}) = 0 \text{ or } f(t) = 0$$

so the mode is at:

$$t = \frac{log(\frac{\beta}{\alpha})}{\beta}$$

as long as $\alpha < \beta$

otherwise the function is decreasing so:

$$t = 0$$

c

$ h(x) = ae^{bx} = e^{log(a) + bx}$ So: $log(h(x)) = log(a) + bx$

```
df_c <- df1 %>%
    filter(between(age, 40, 100)) %>%
    mutate(loghx = log(hx))

df_1961 <- df_c %>% filter(year == 1961)
df_2011 <- df_c %>% filter(year == 2011)

lm1961 <- lm(loghx ~ age, data = df_1961)
lm2011 <- lm(loghx ~ age, data = df_2011)

coef1961 <- coef(lm1961)
coef2011 <- coef(lm2011)

a1961 = unname(exp(coef1961[1]))
a2011 = unname(exp(coef2011[1]))

b1961 = unname(coef1961[2])
b2011 = unname(coef2011[2])
```

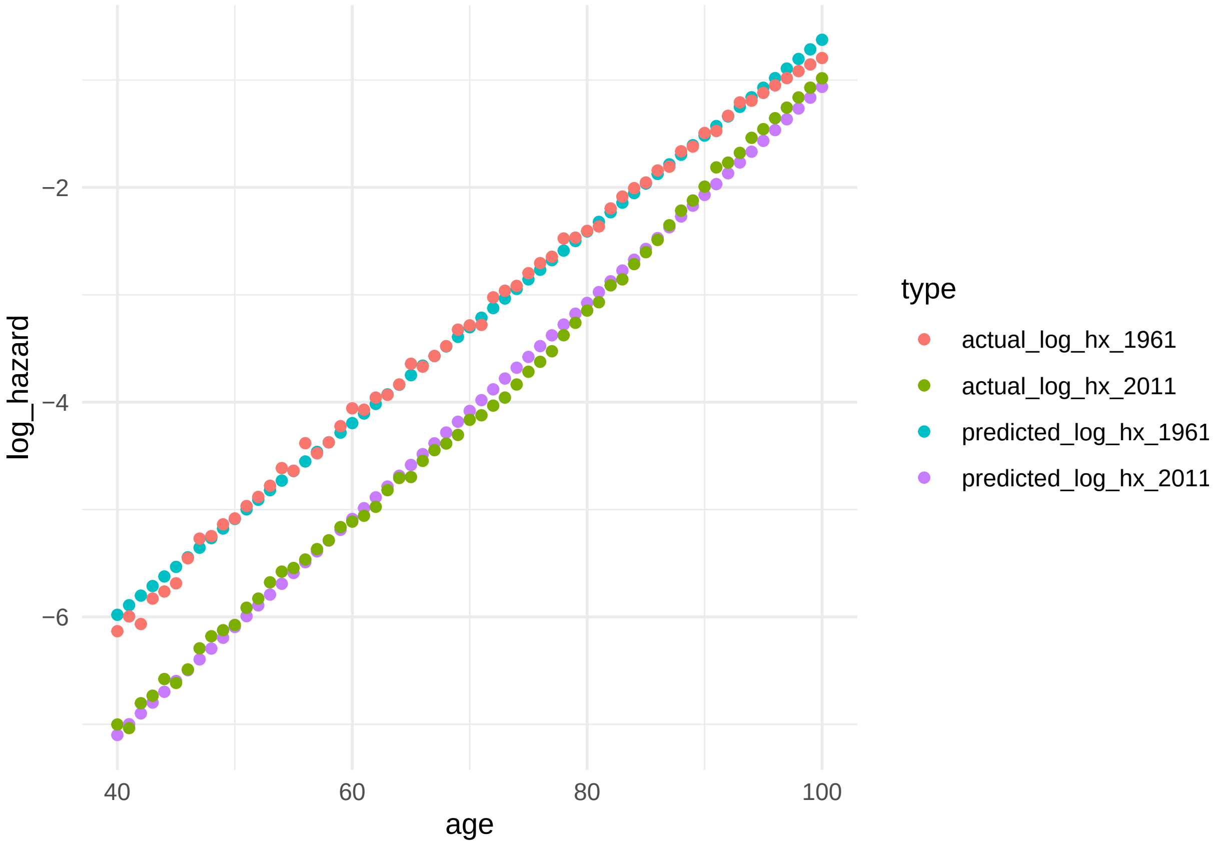
The values for 1961 are: alpha of 7.1168697×10^{-5} and beta of 0.0892529 compared to the values for 2011 : alpha of 1.4755769×10^{-5} and beta of 0.1006012

The meaning of alpha is the starting level of mortality (much higher for 1961) and beta gives the increase in mortality over time which surprisingly is higher for 2011. Perhaps lower infant mortality screws with us a tiny bit and makes it seem like people die faster with age in 2011 than they did in 1961 just because so many of them already died before 40 where we start.

d

```
preds_d <- tibble(age = seq(from = 40, to = 100, by = 1))
preds_d$predicted_log_hx_1961 <- predict(lm1961, newdata = preds_d)
preds_d$predicted_log_hx_2011 <- predict(lm2011, newdata = preds_d)
preds_d$actual_log_hx_1961 <- df_1961$loghx
```

```r
preds_d$actual_log_hx_2011 <- df_2011$loghx

preds_d_long <- preds_d %>% pivot_longer(cols = c("predicted_log_hx_1961", "predicted_log_hx_2011", "ac

preds_d_long %>%
    ggplot(aes(x = age, y = log_hazard, color = type)) +
    geom_point() +
    theme_minimal()
```

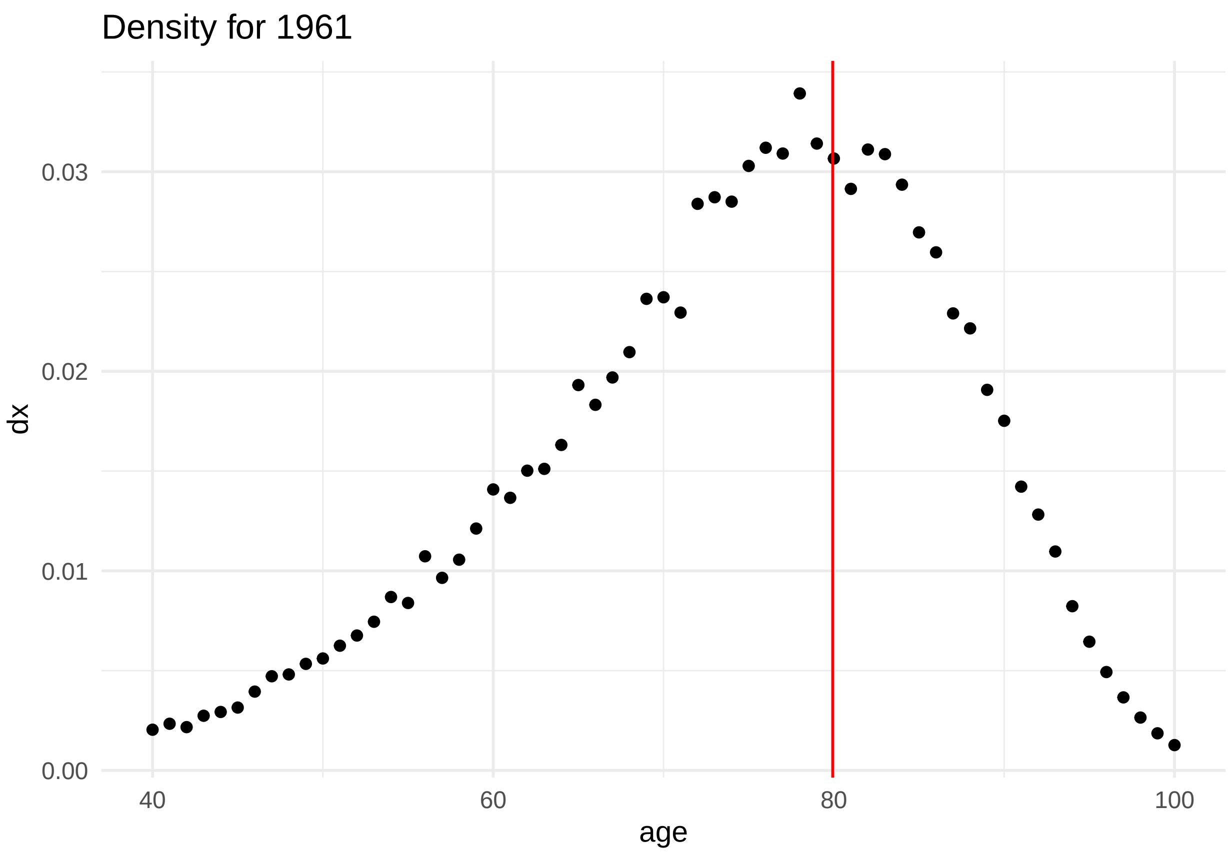


They both seem to fit surprisingly well. There are some minor patterns in the predicted vs actual for 2011 between the ages of 70 and 80 where the actual log hazard seems to be lower, and later on when actual log hazard seems to be higher for those 90+. For 1961 model it seems to be the opposite for the super old - we overestimate the log hazard for those pushing 100. Overall I would say the assumption is quite reasonable.

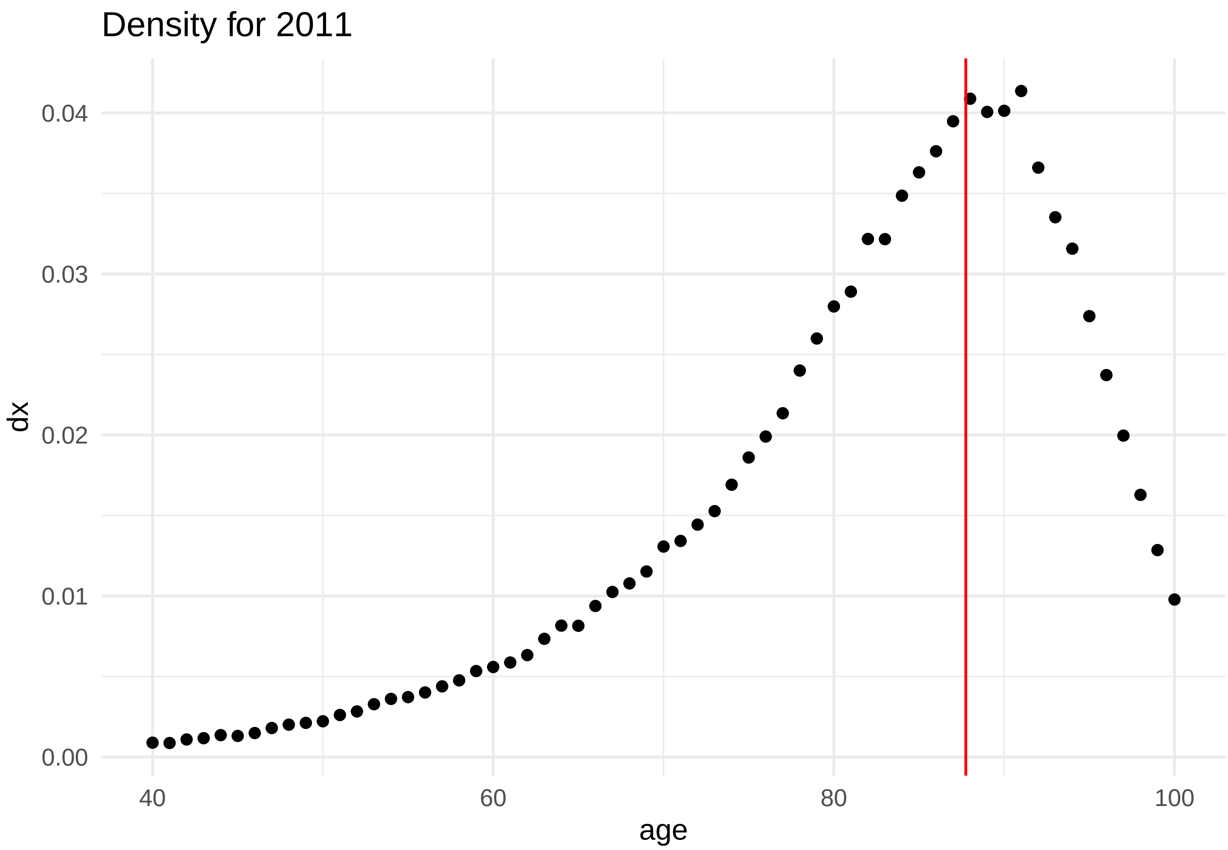
e

```r
mode1961 <- log(b1961 / a1961) / b1961
mode2011 <- log(b2011 / a2011) / b2011

df_e <- df_c %>% filter(year %in% c(2011, 1961))

df_e %>% filter(year == 1961) %>%
    ggplot(aes(x = age, y = dx)) +
    geom_point() +
    geom_vline(xintercept = mode1961, color = "red") +
    theme_minimal() +
```

```
    labs(title = "Density for 1961")
```



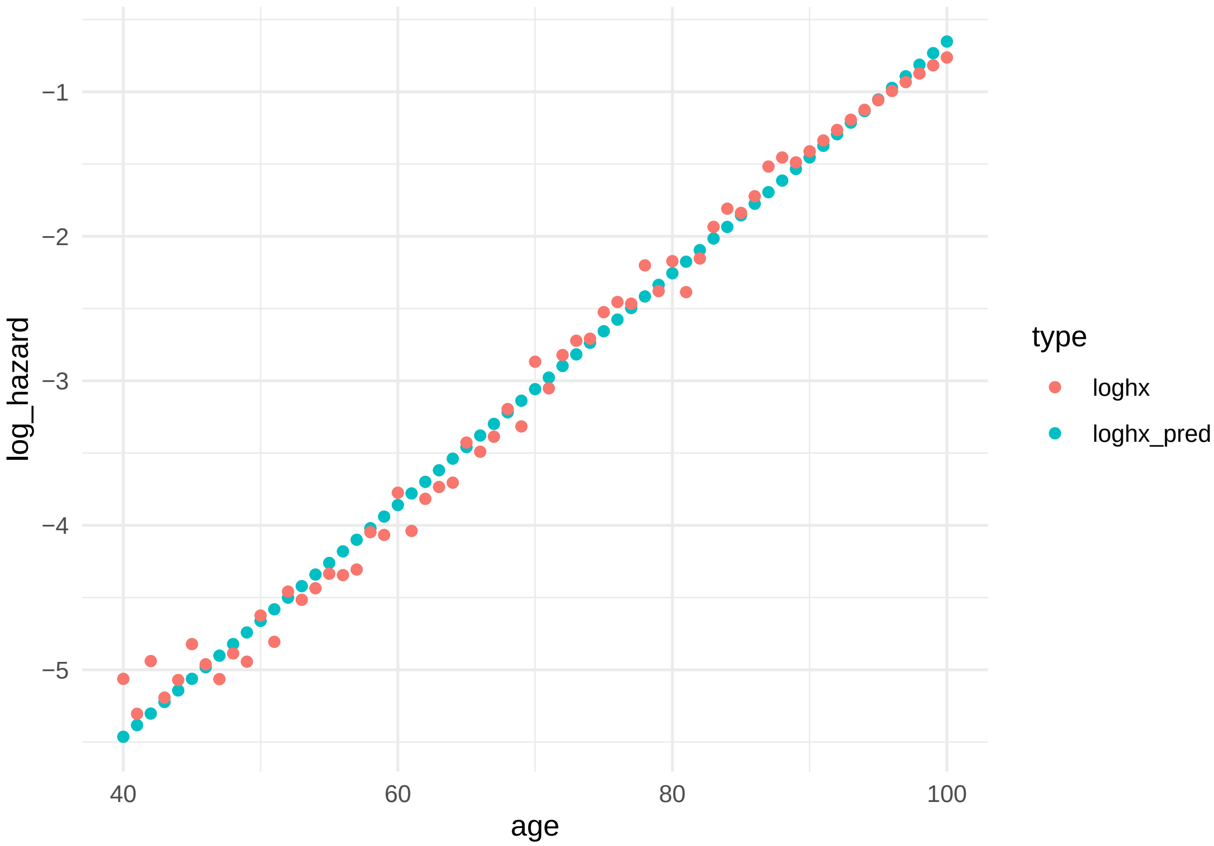

```
df_e %>% filter(year == 2011) %>%
    ggplot(aes(x = age, y = dx)) +
    geom_point() +
    geom_vline(xintercept = mode2011, color = "red") +
    theme_minimal() +
    labs(title = "Density for 2011")
```



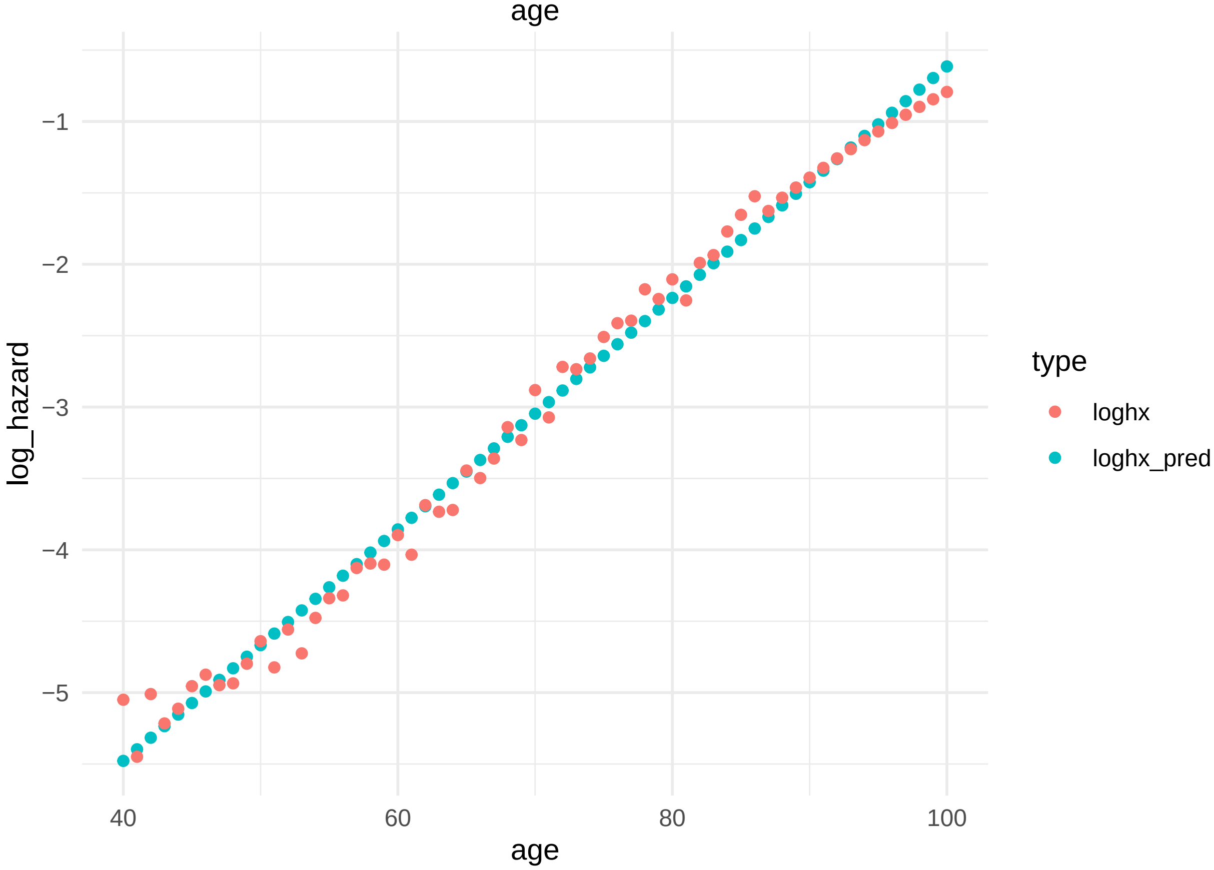
f

I could probably figure this out with map(lm) but I don't wanna right now.

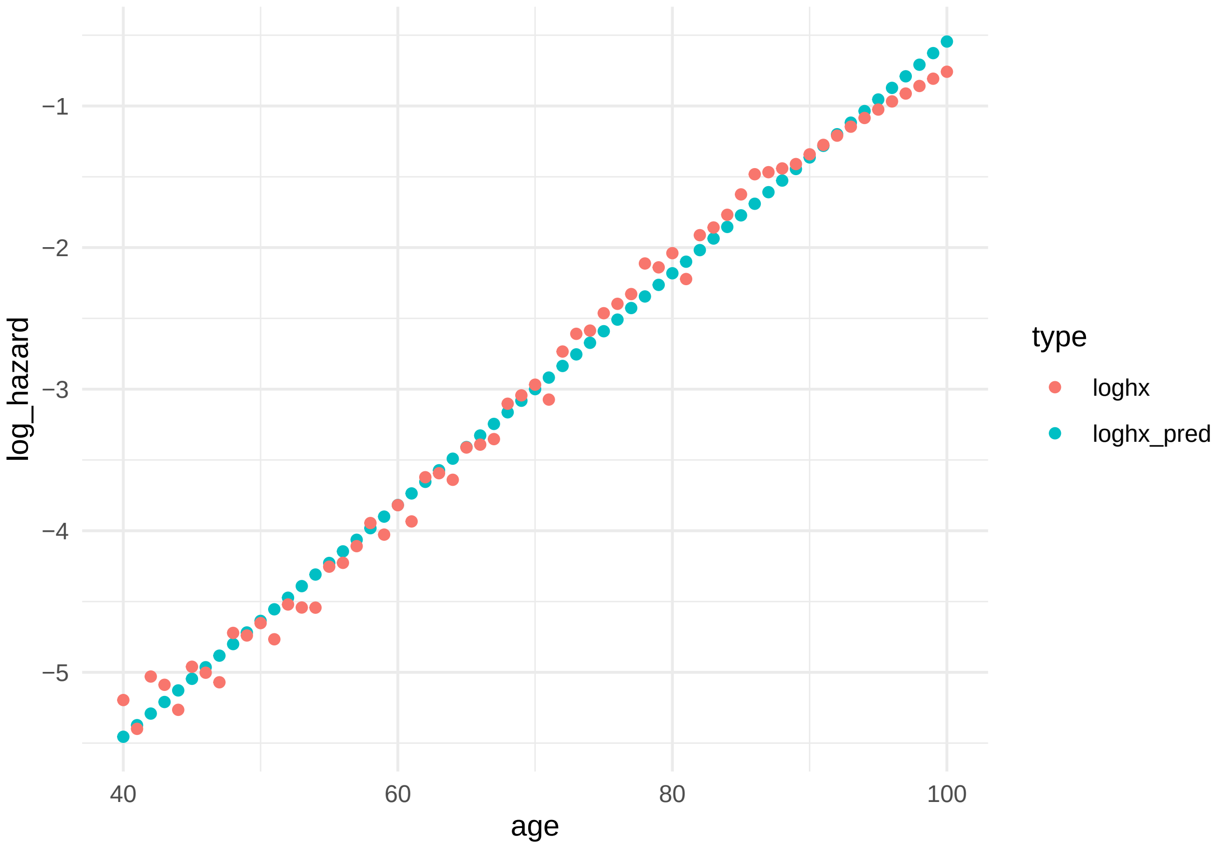
```
years <- unique(df_c$year)
alphas <- rep(NA, length(years))
betas <- rep(NA, length(years))


for (i in 1:length(years)) {
     # fit model
    df_model <- df_c %>%
        filter(year == years[i])
    lm_loop <- lm(loghx ~ age, data = df_model)
    coef_model <- coef(lm_loop)

    alphas[i] <- unname(exp(coef_model[1]))
    betas[i] <- unname(coef_model[2])

    df_model$loghx_pred <- predict(lm_loop, newdata = df_model)

    df_plot <- df_model %>% pivot_longer(cols = c("loghx_pred", "loghx"), names_to = "type", values_to =
    p <- df_plot %>% ggplot(aes(x = age, y = log_hazard, color = type)) +
    geom_point() +
    theme_minimal()
    print(p)
}
```



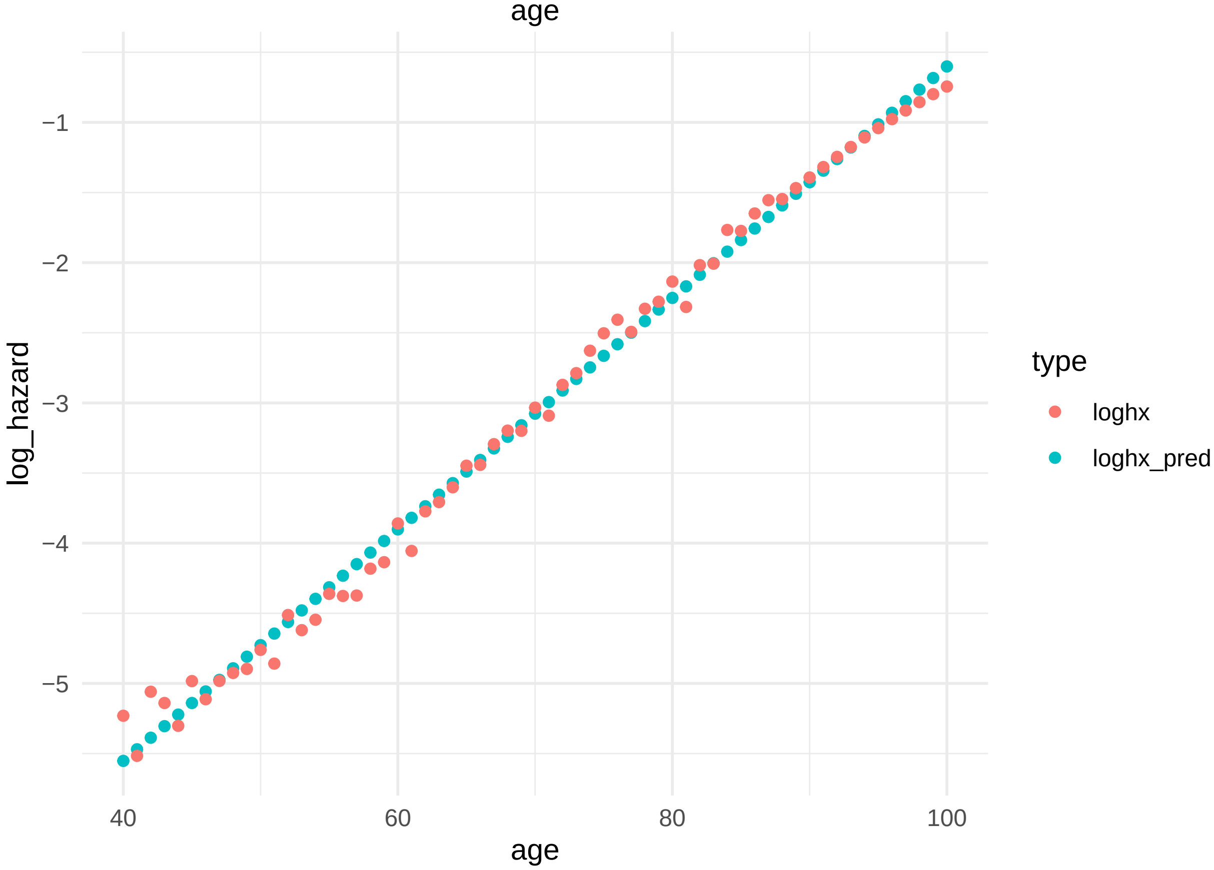
−1
−2
−3
−4
−5
log_hazard
40
60
80
100
age
type
loghx
loghx_pred



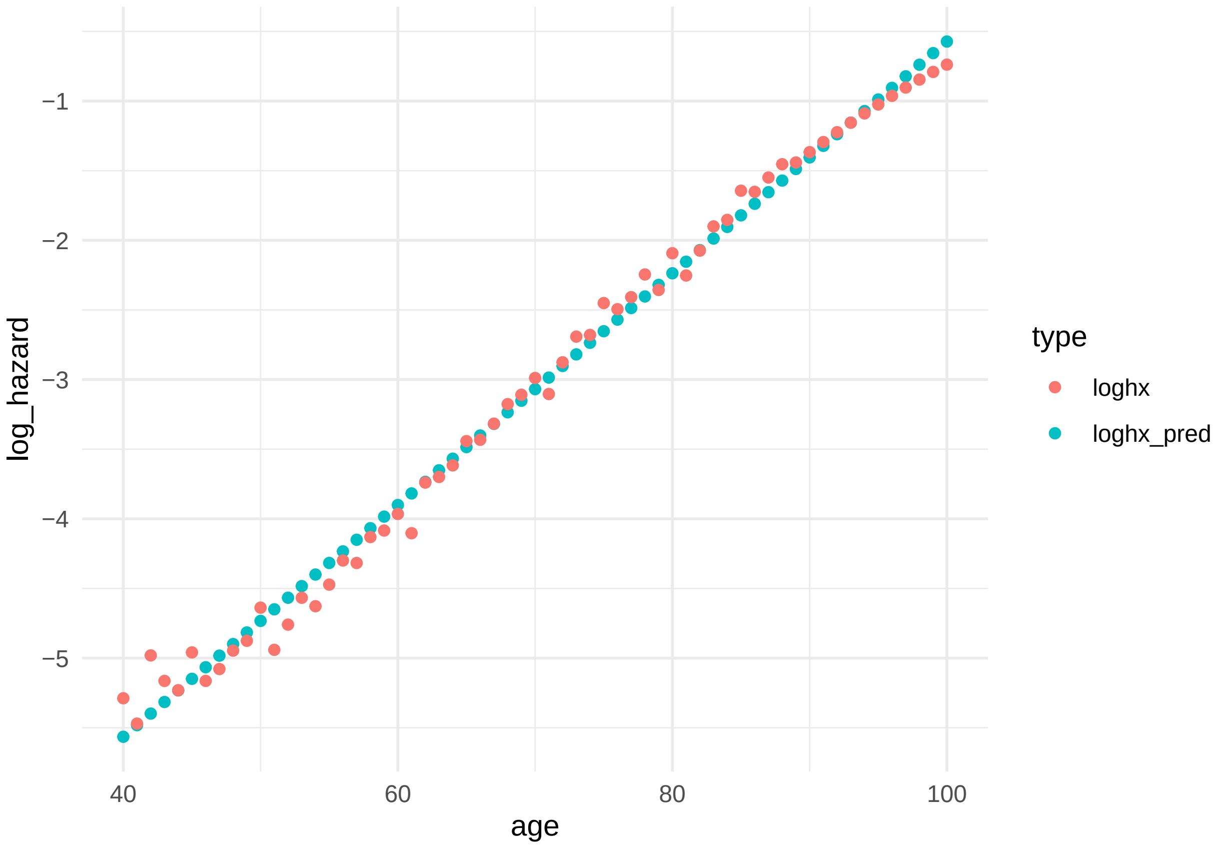
−1
−2
−3
−4
−5
log_hazard
40
60
80
100
age
type
loghx
loghx_pred



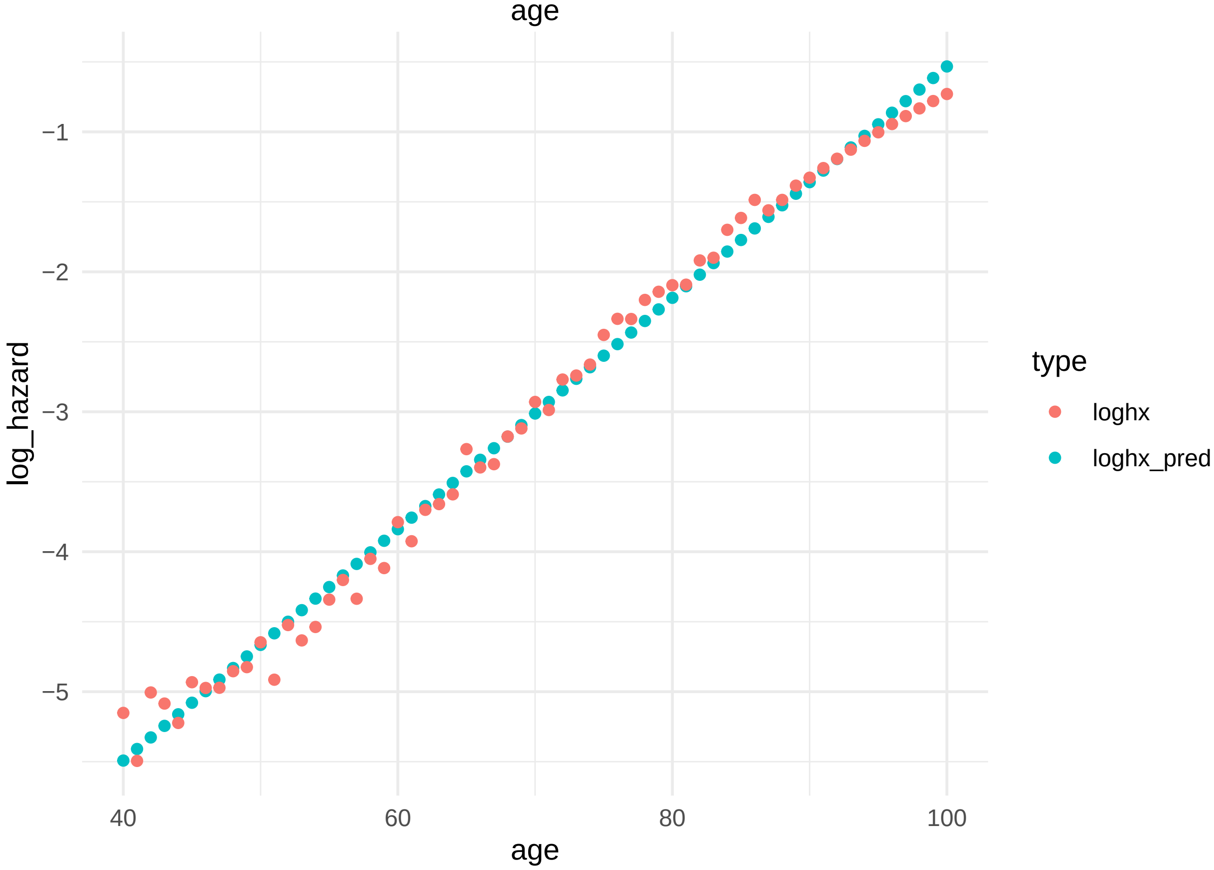
−1
−2
−3
−4
−5
log_hazard
40
60
80
100
age
type
loghx
loghx_pred



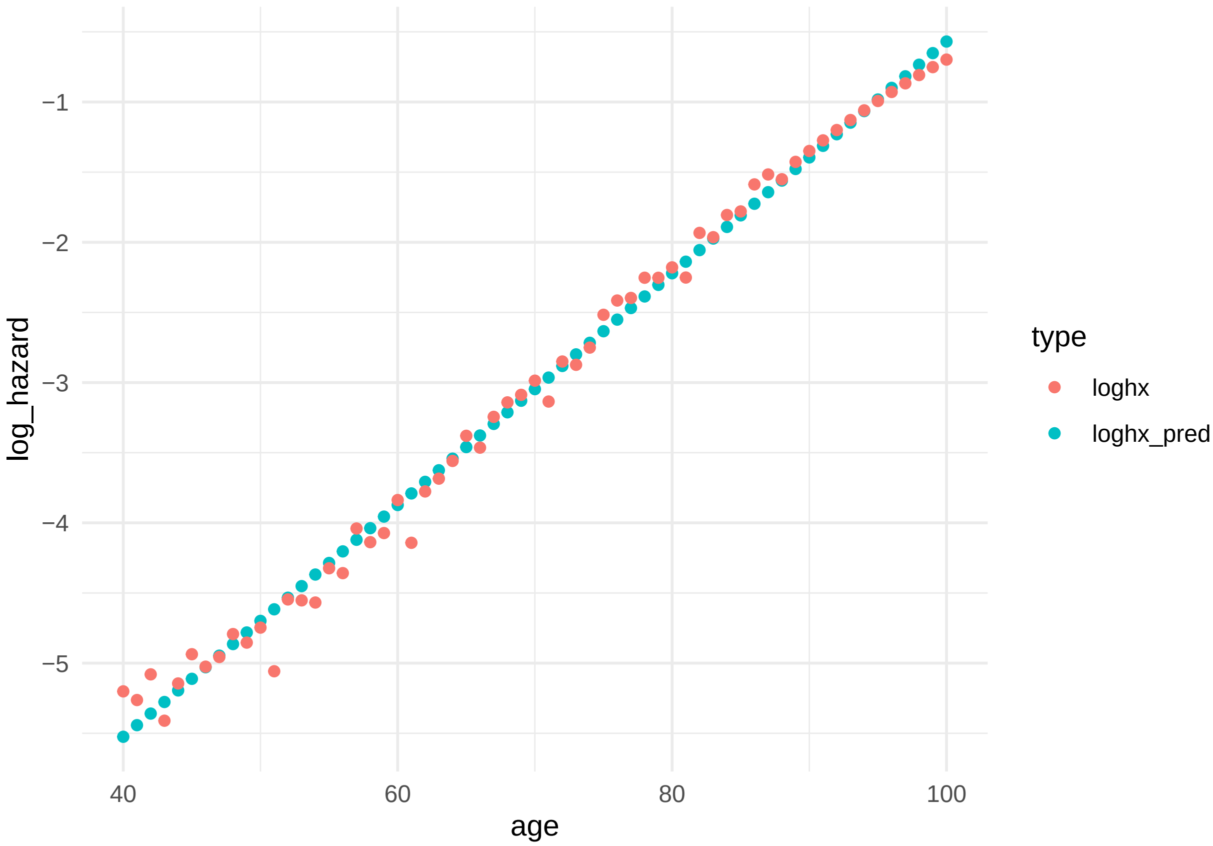
−1
−2
−3
−4
−5
log_hazard
40
60
80
100
age
type
loghx
loghx_pred

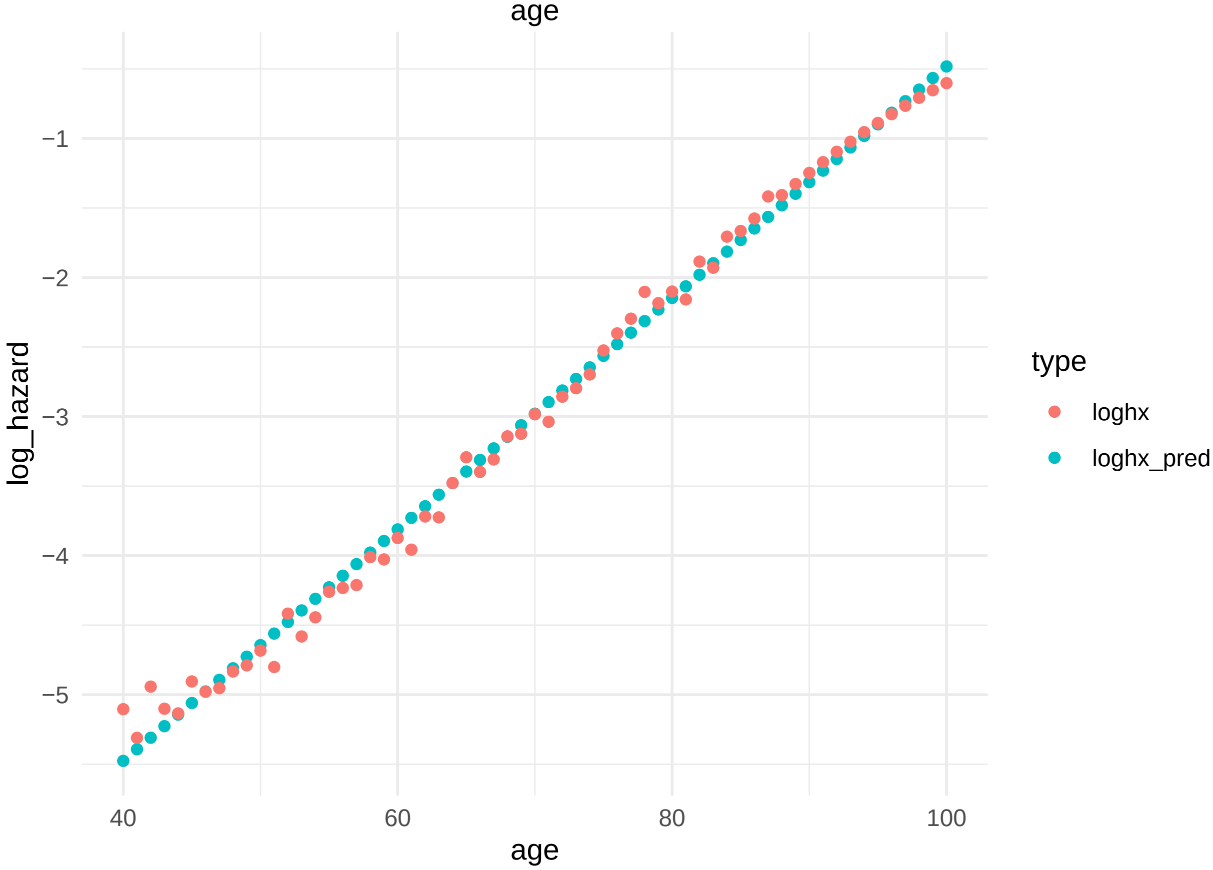


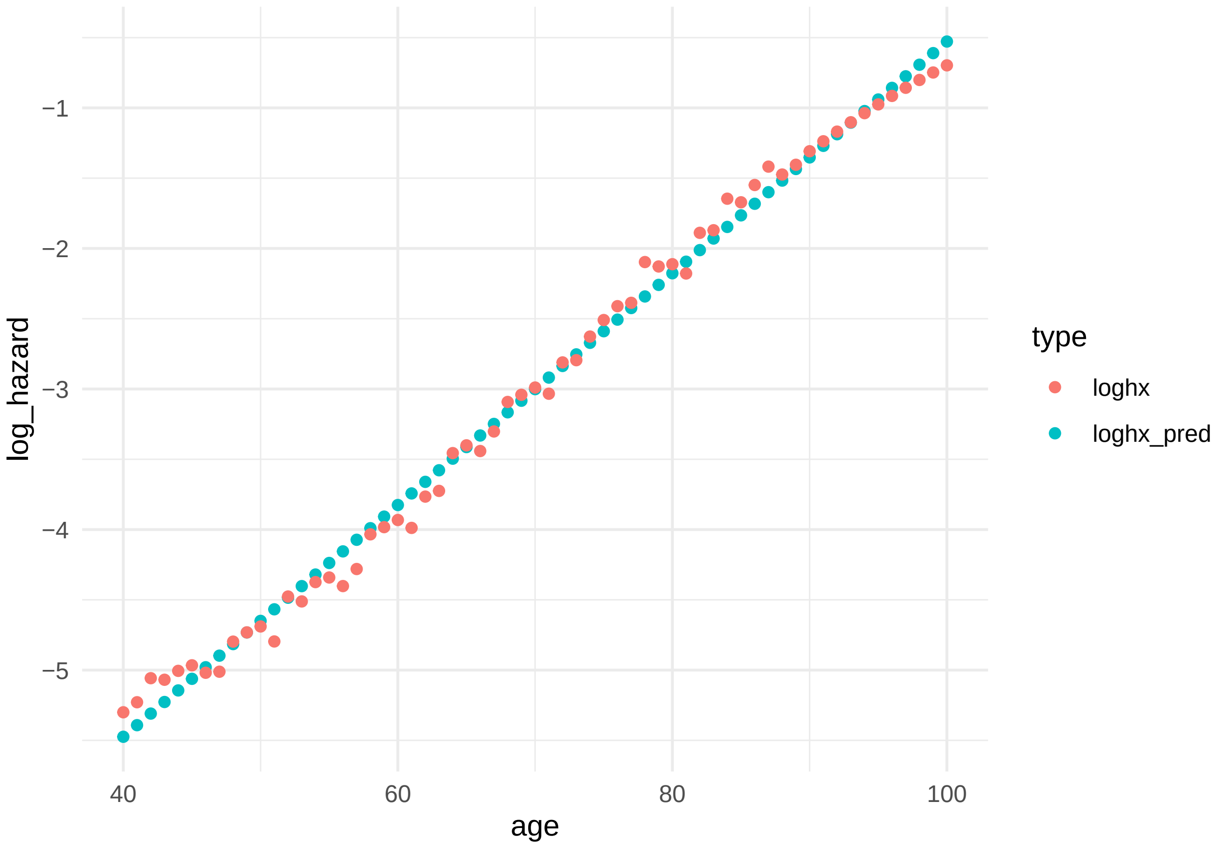
log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred



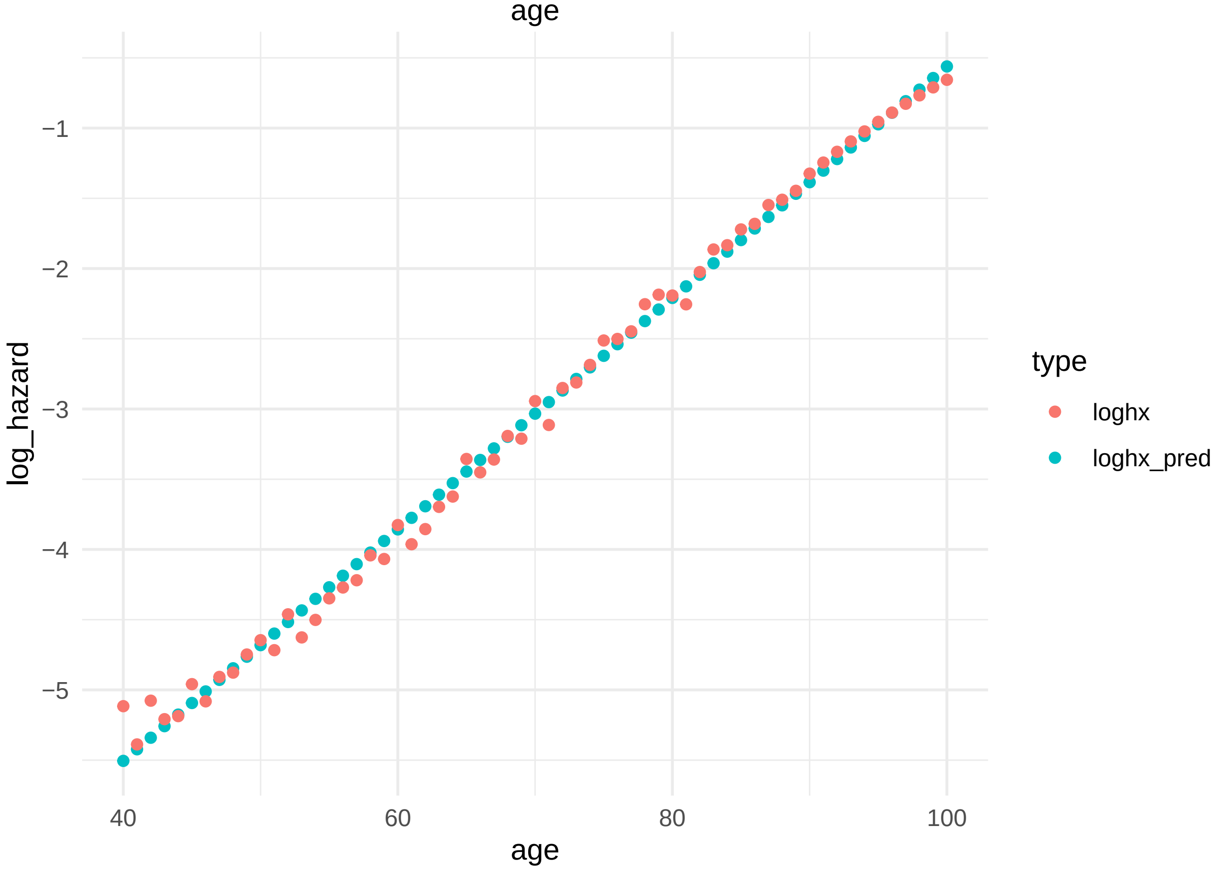
log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
age
type
loghx
loghx_pred

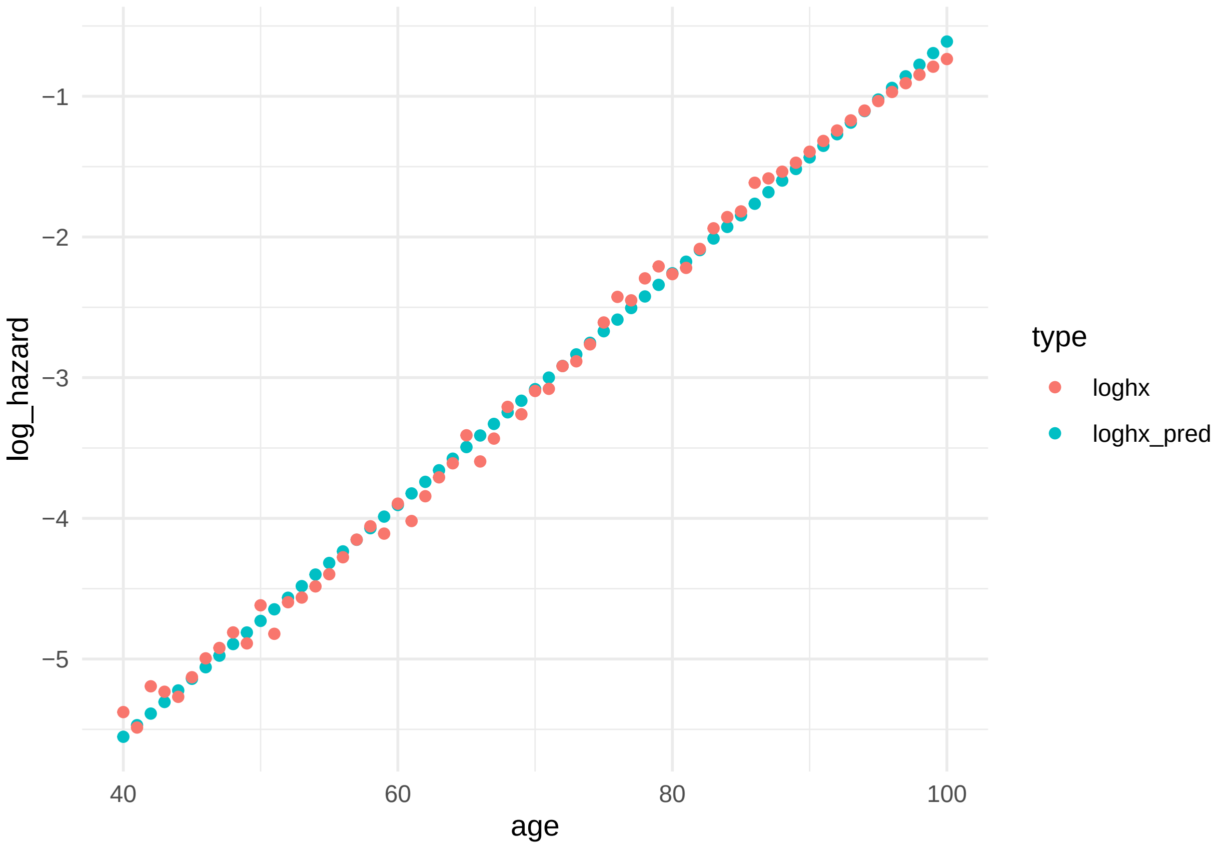

log_hazard
age
type
loghx
loghx_pred

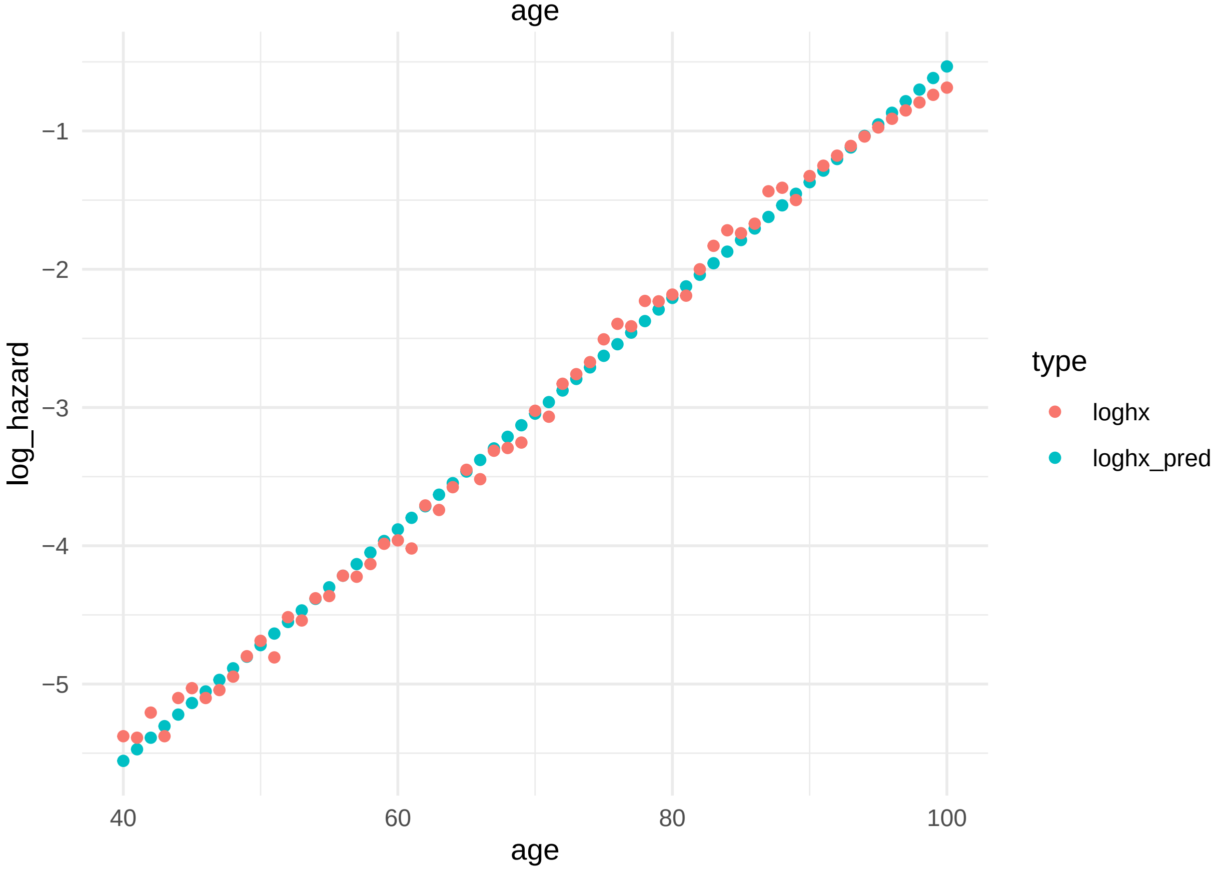


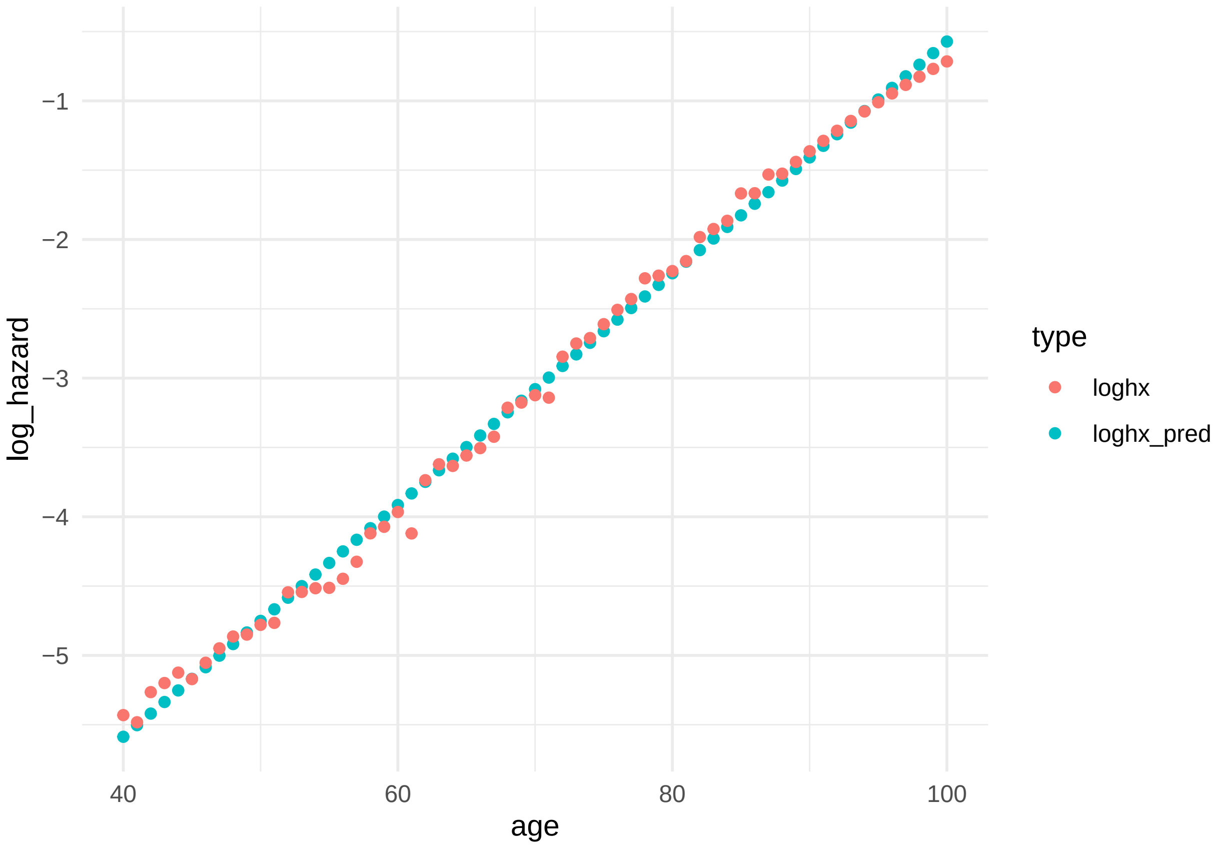
log_hazard
age
type
loghx
loghx_pred



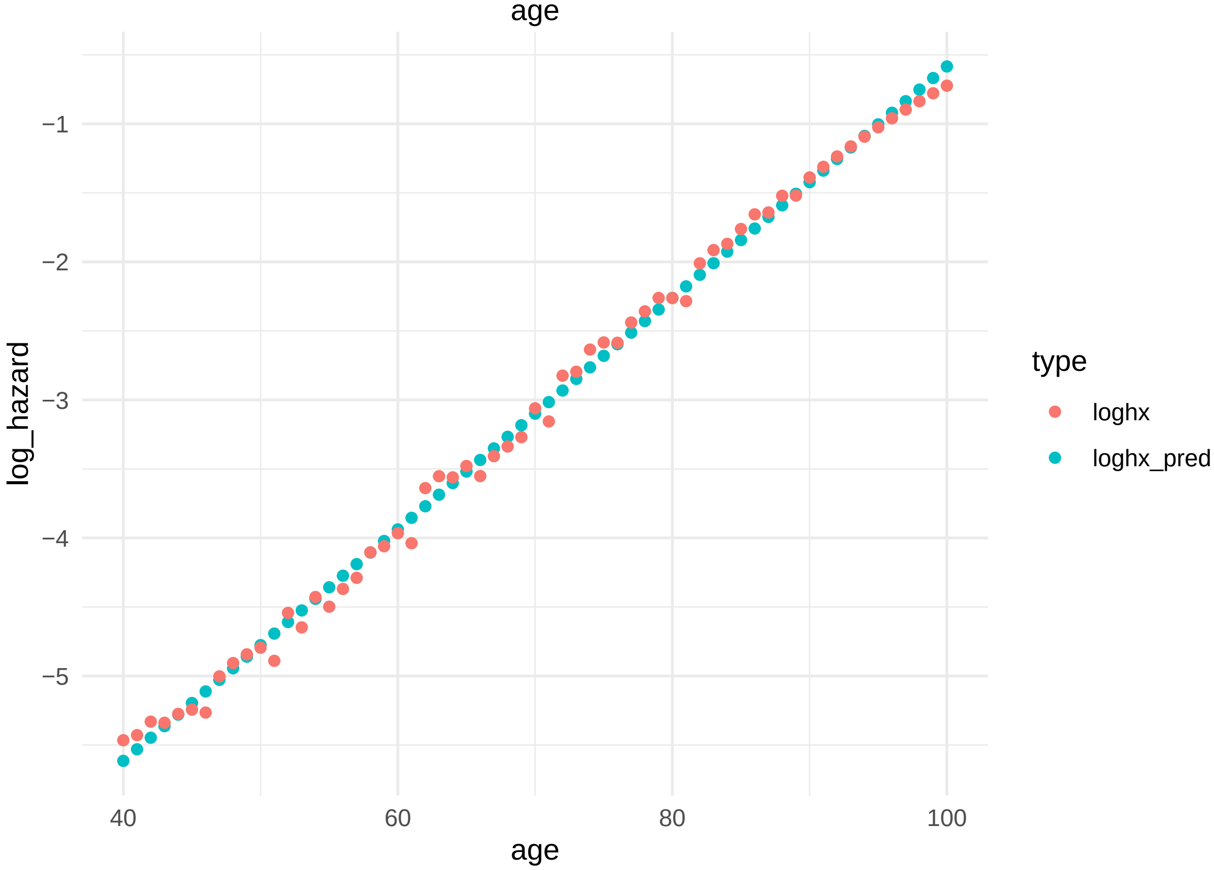
log_hazard
age
type
loghx
loghx_pred


log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred

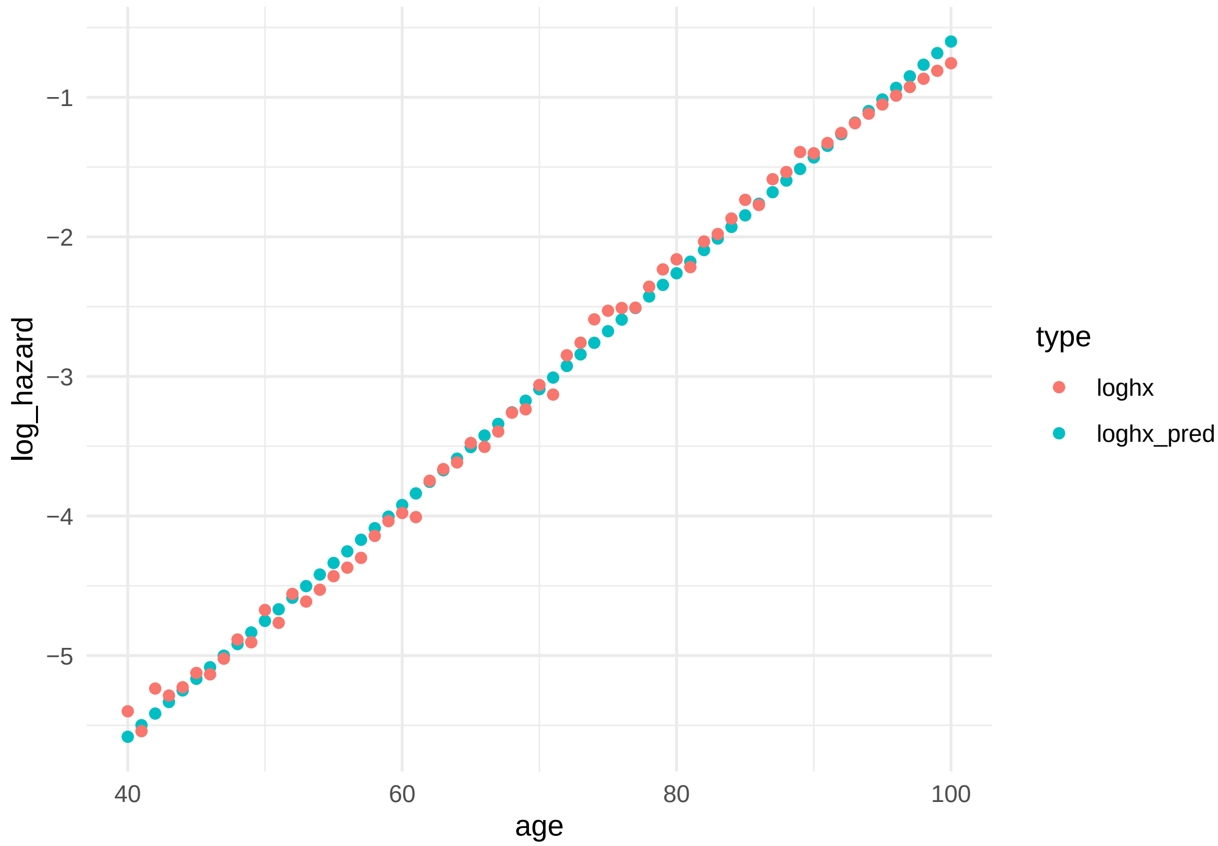

log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred

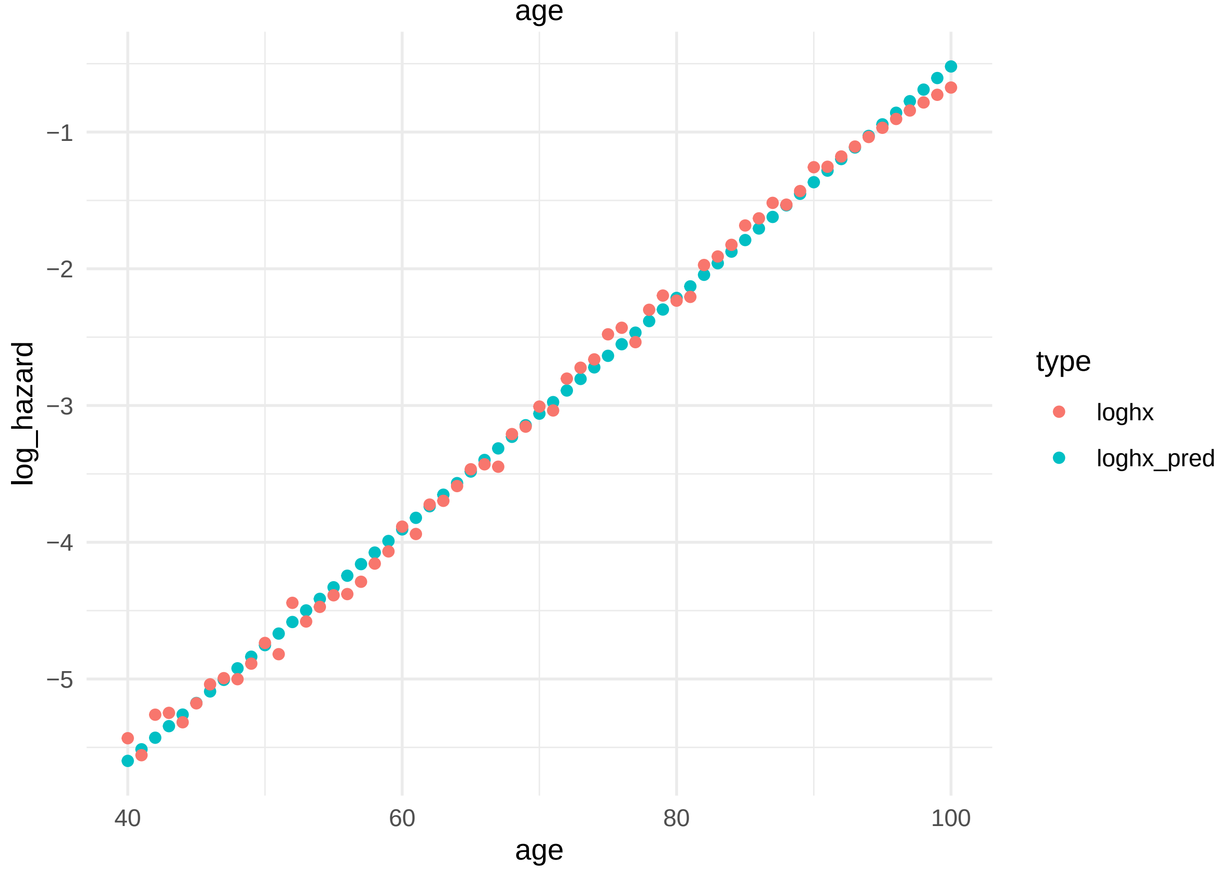


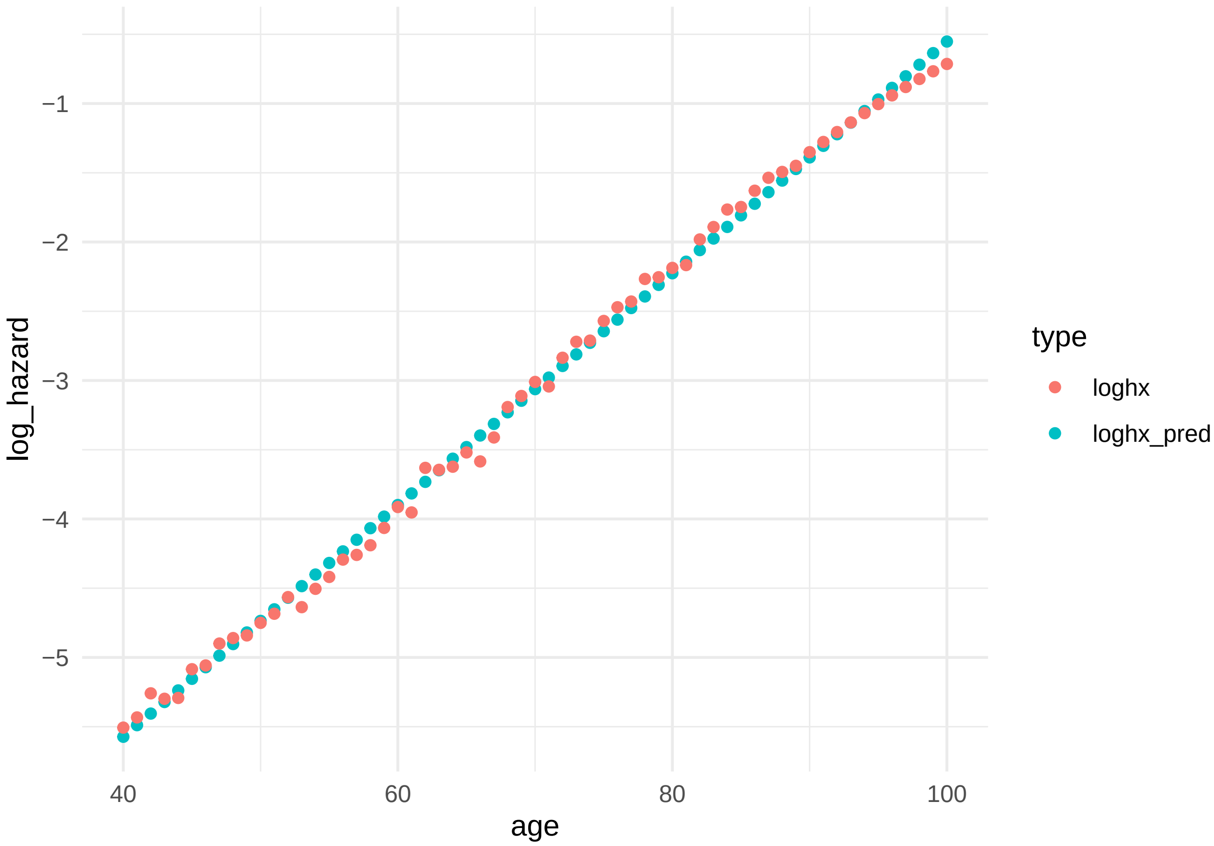
log_hazard
−1
−2
−3
−4
−5
age
40
60
80
100
type
loghx
loghx_pred

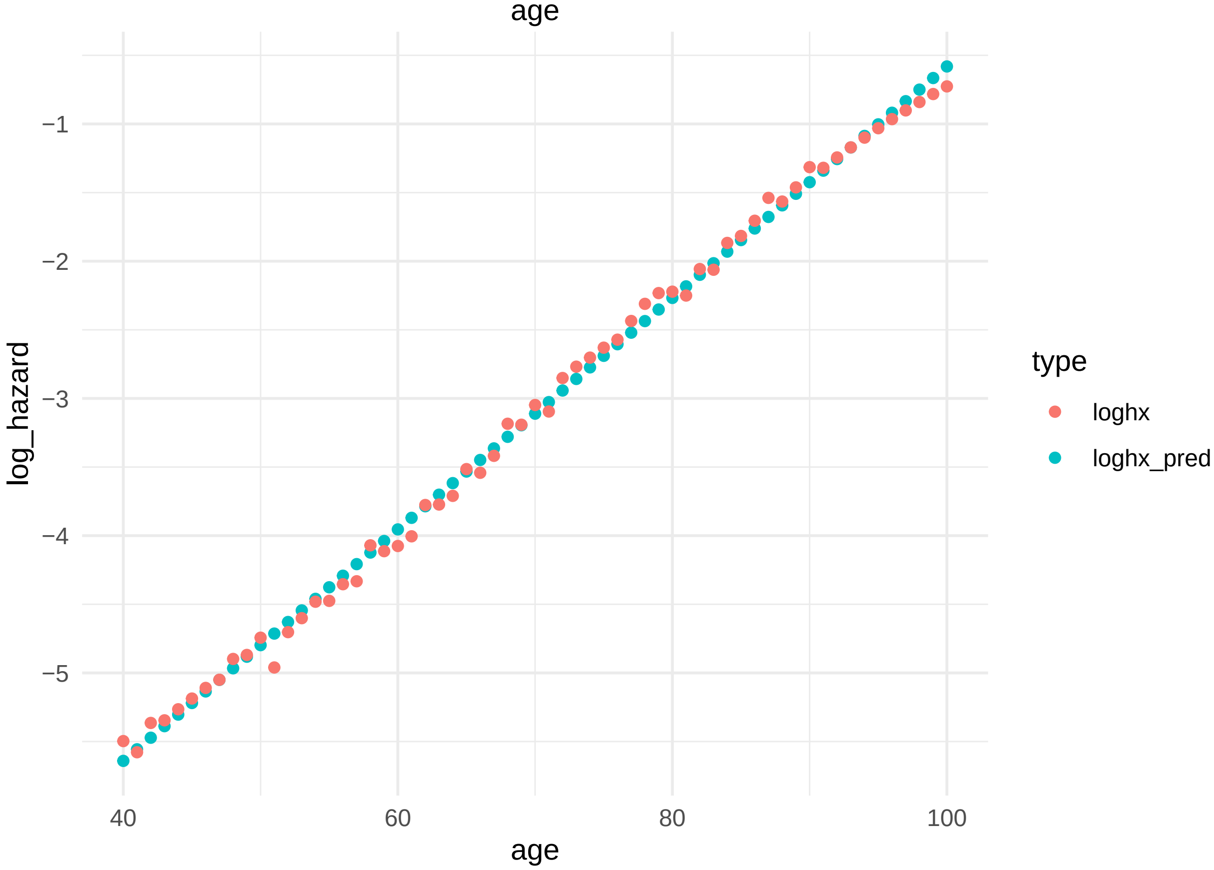


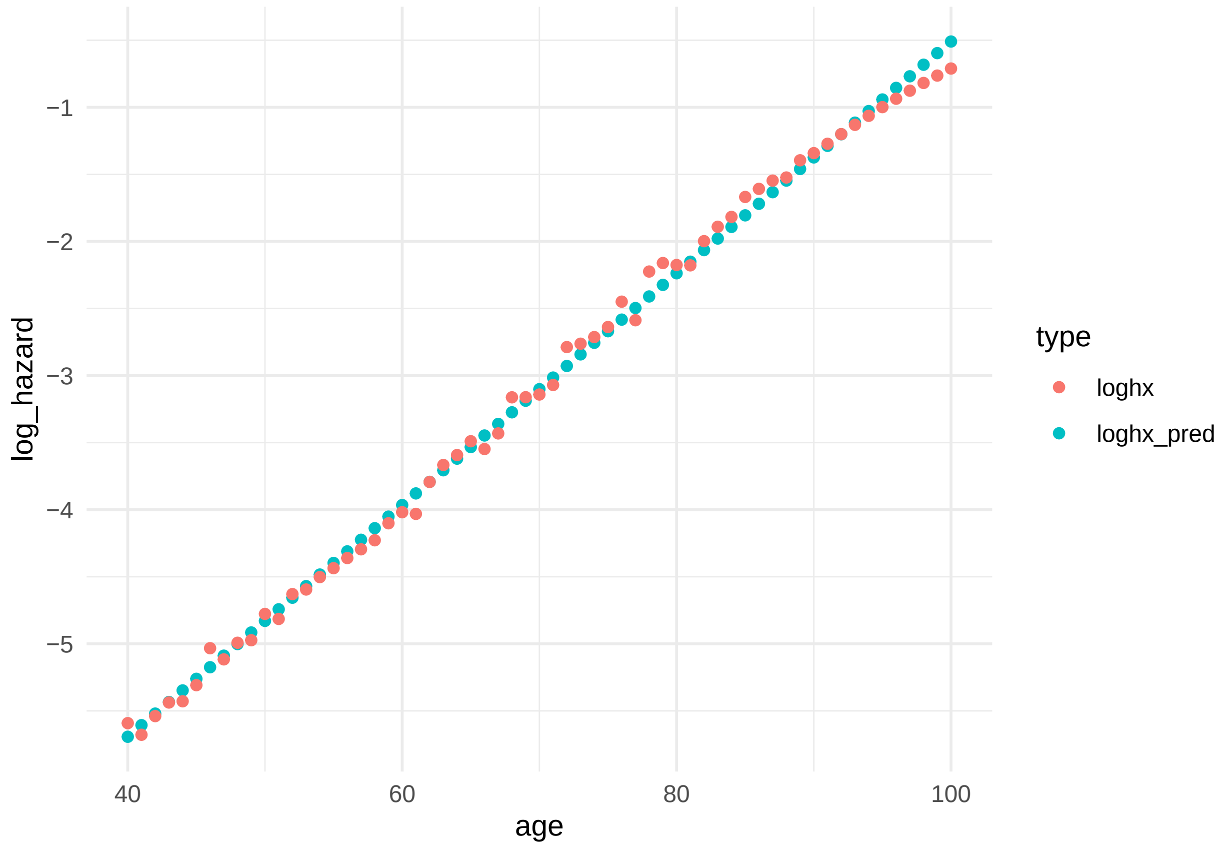
log_hazard
−1
−2
−3
−4
−5
age
40
60
80
100
type
loghx
loghx_pred

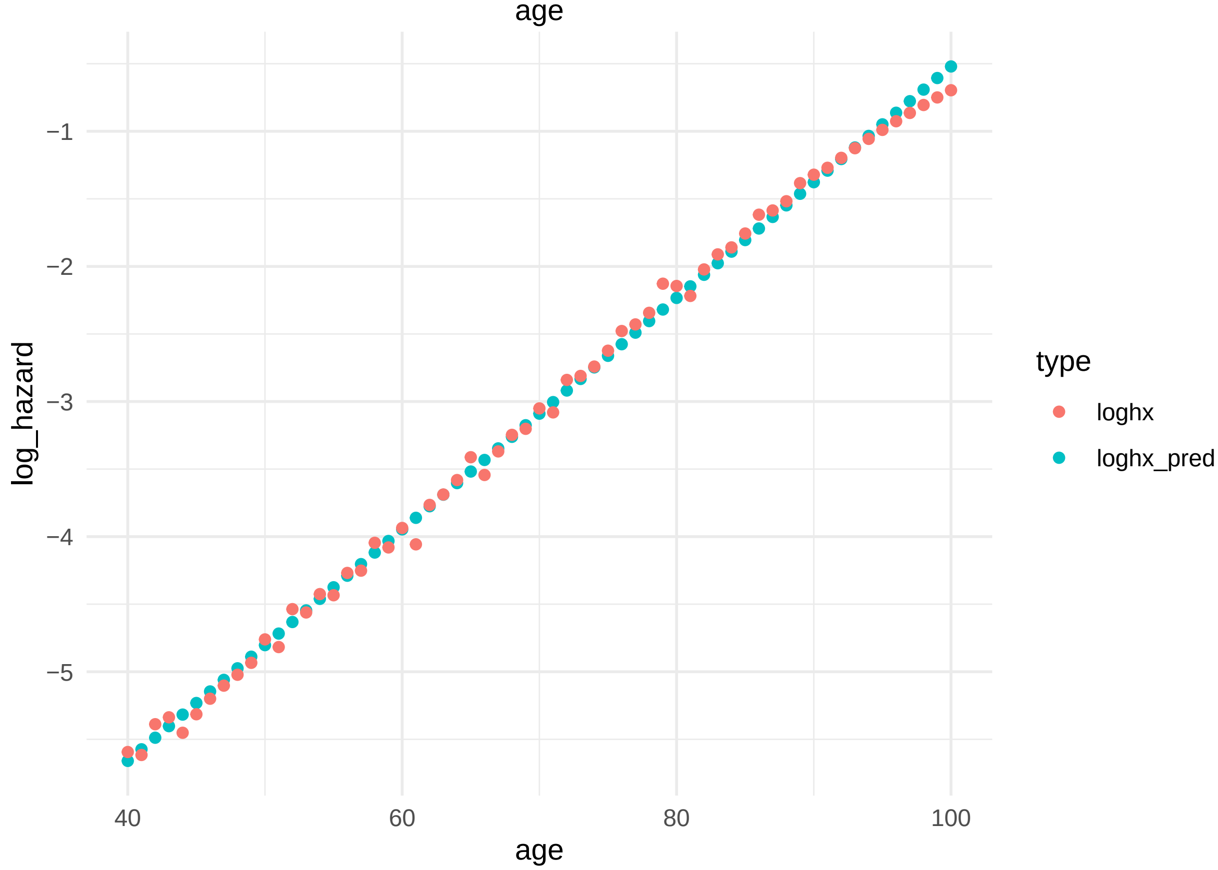

log_hazard
age
type
loghx
loghx_pred

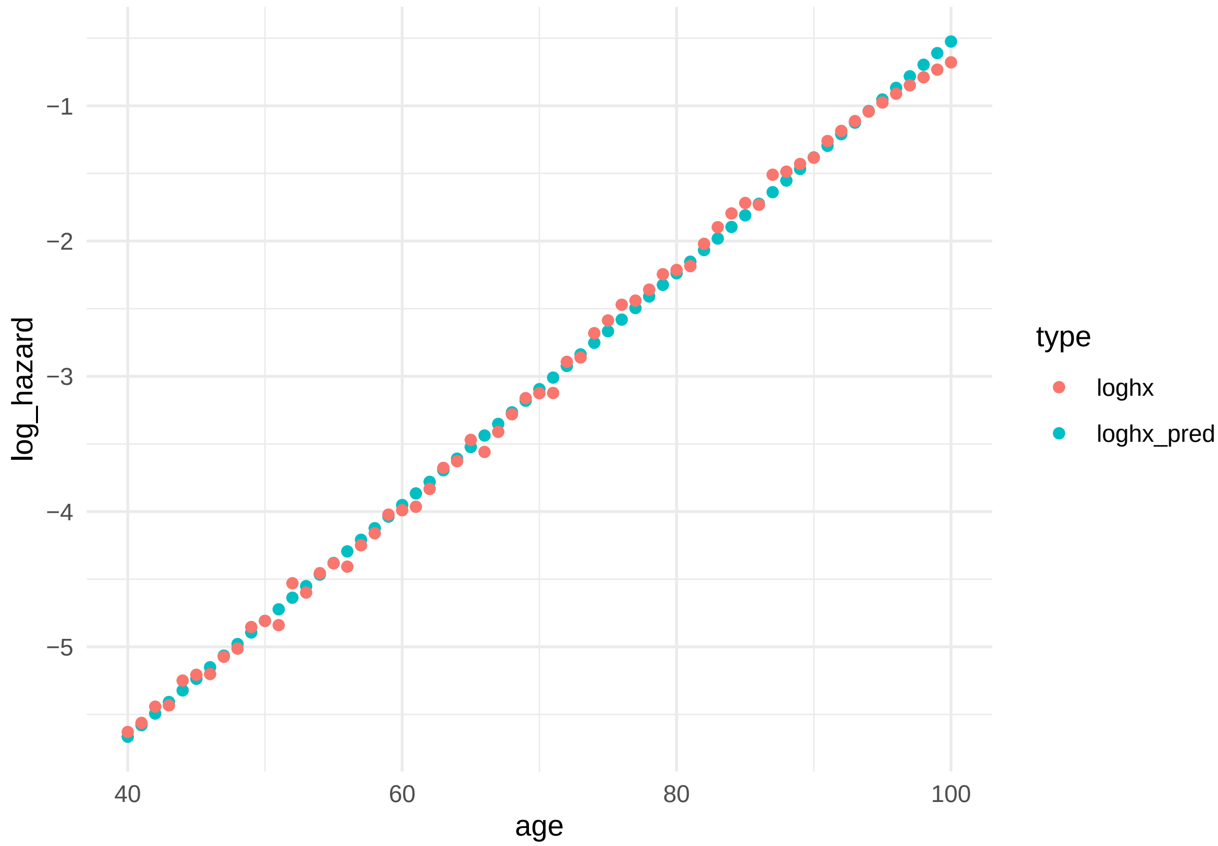

log_hazard
age
type
loghx
loghx_pred


log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred

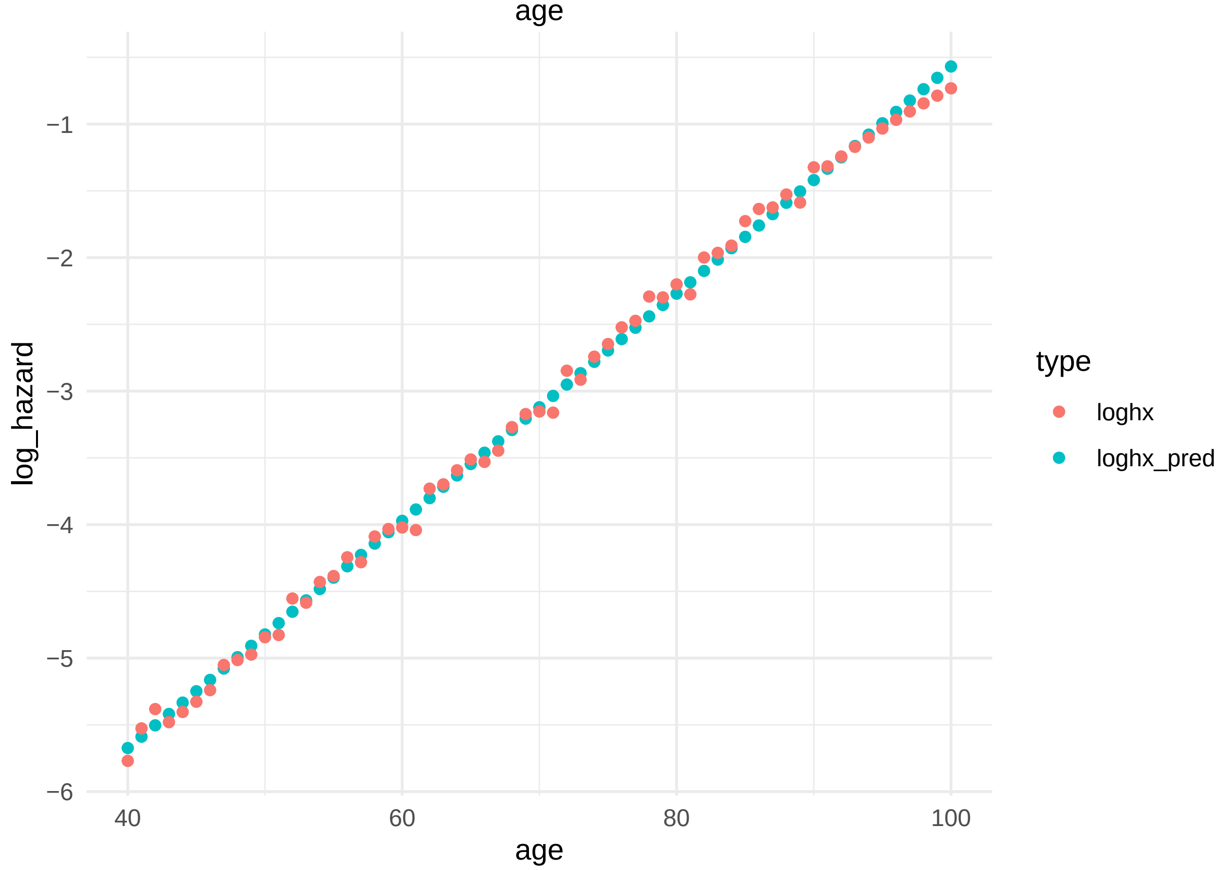

log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred

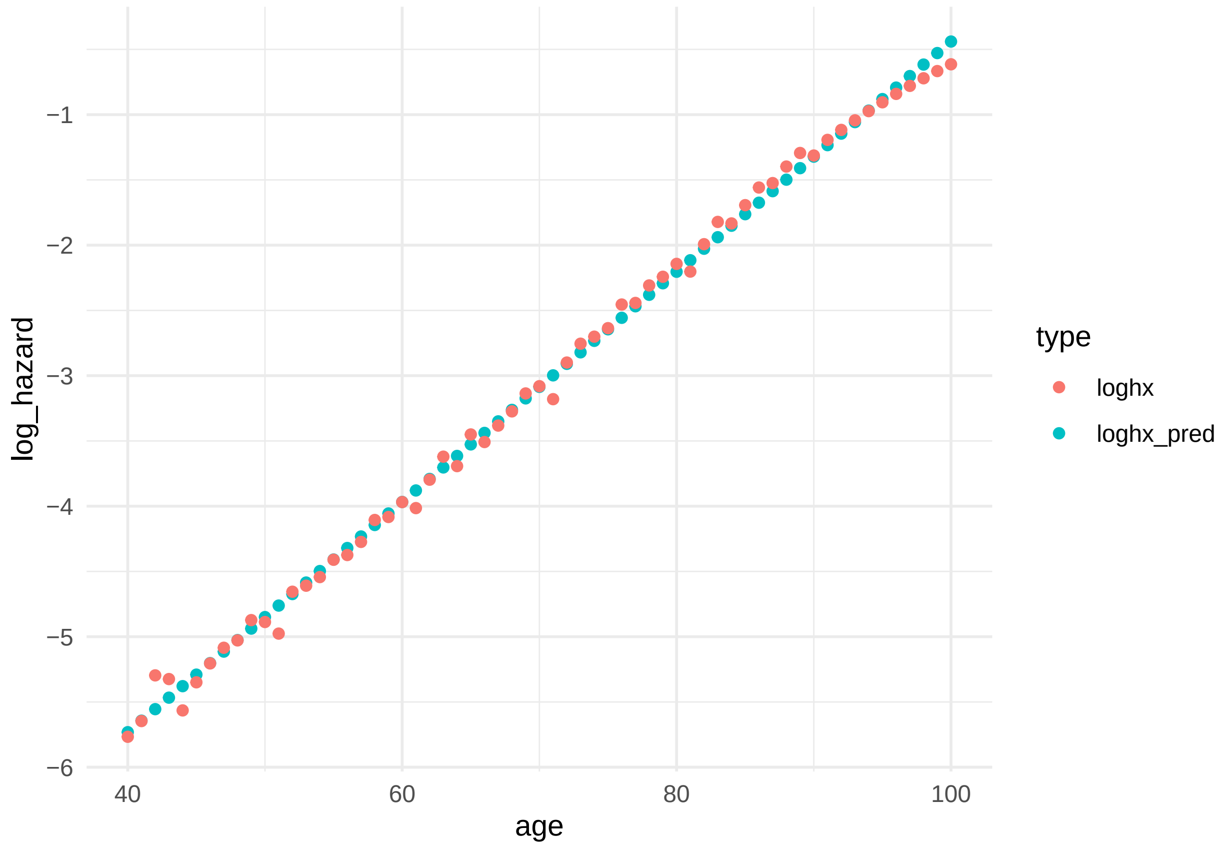

log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred



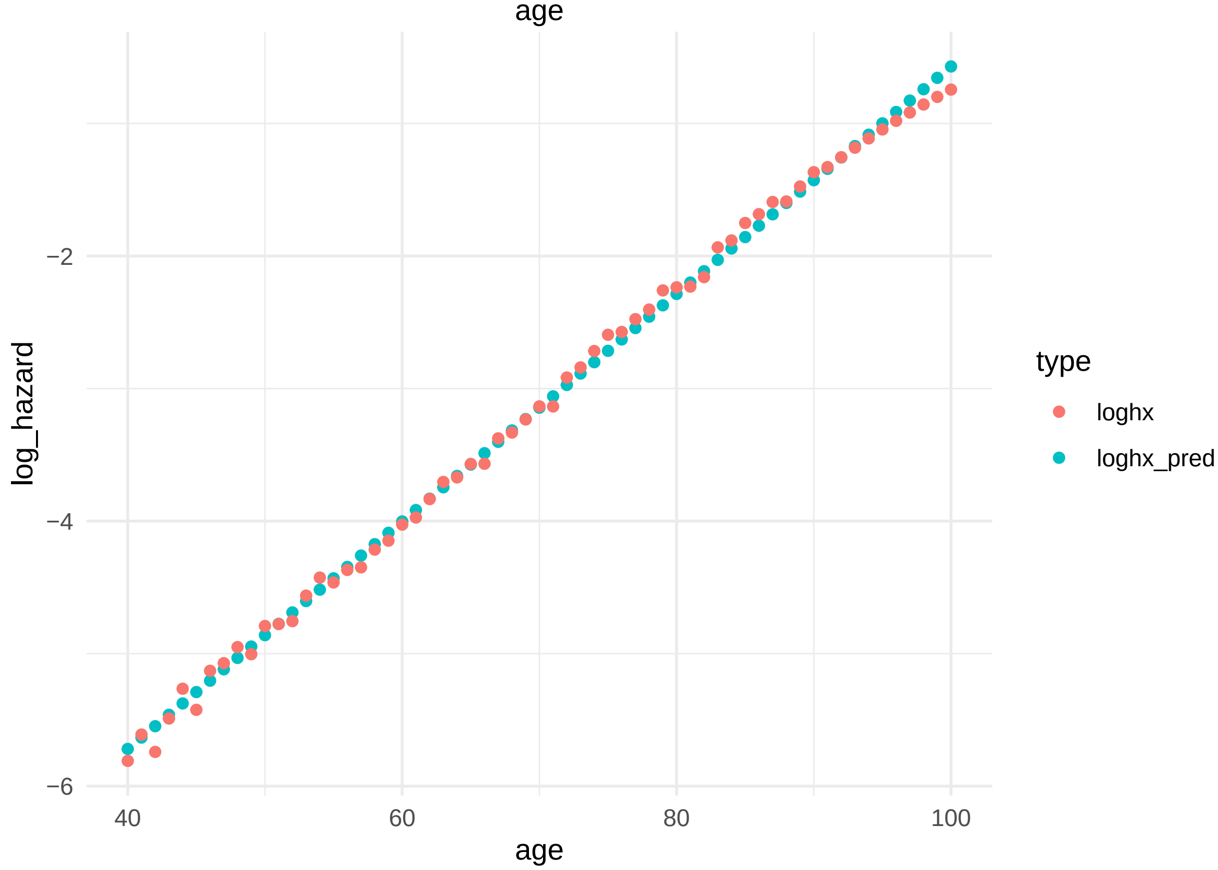
log_hazard
−1
−2
−3
−4
−5
40
60
80
100
age
type
loghx
loghx_pred



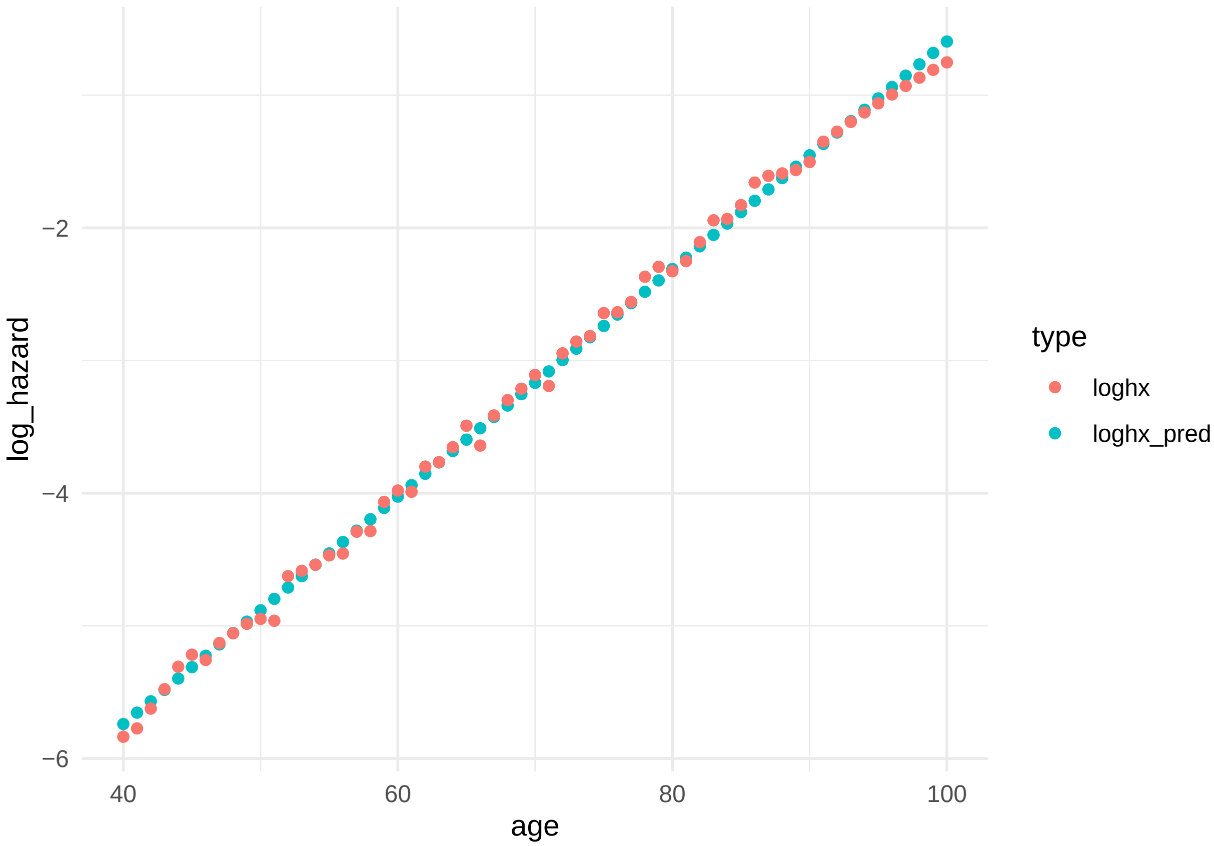
log_hazard
−1
−2
−3
−4
−5
−6
40
60
80
100
age
type
loghx
loghx_pred



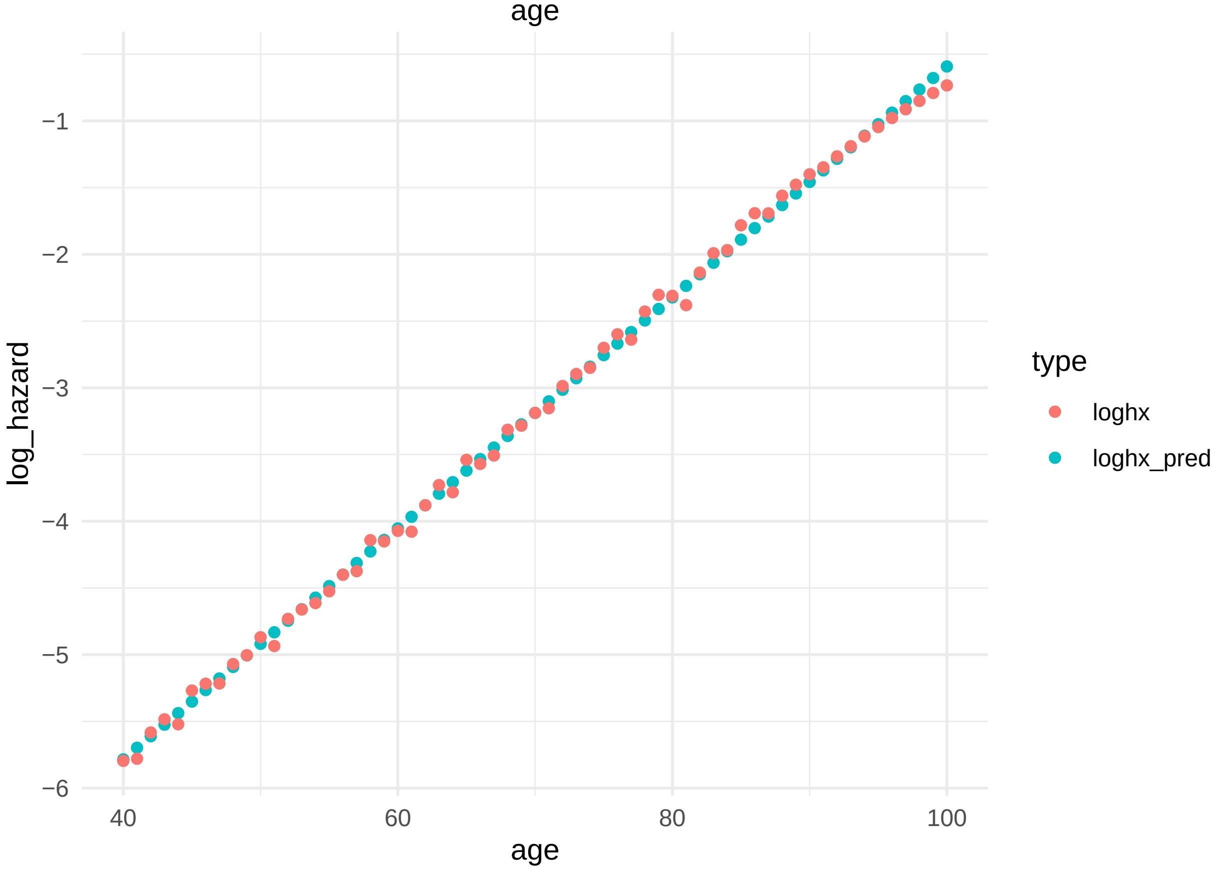
log_hazard
−1
−2
−3
−4
−5
−6
40
60
80
100
age
type
loghx
loghx_pred



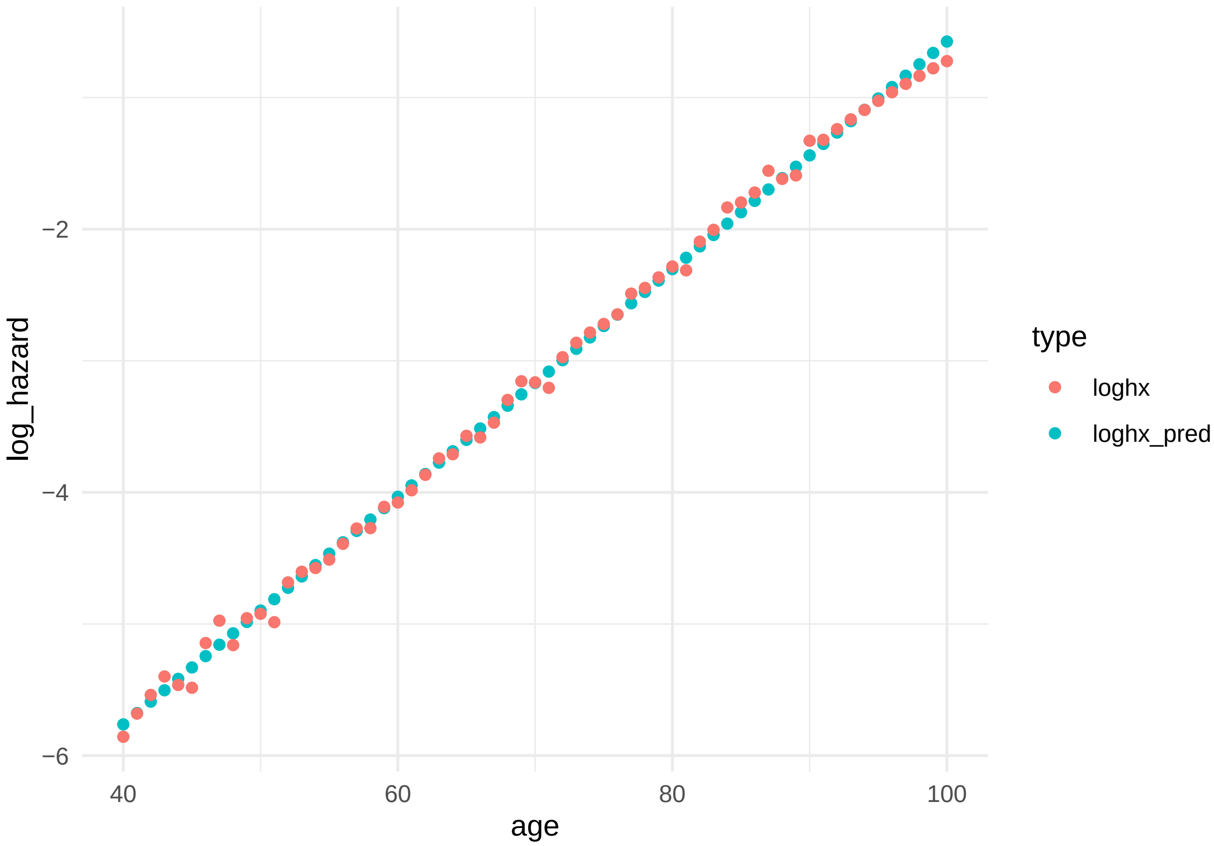
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



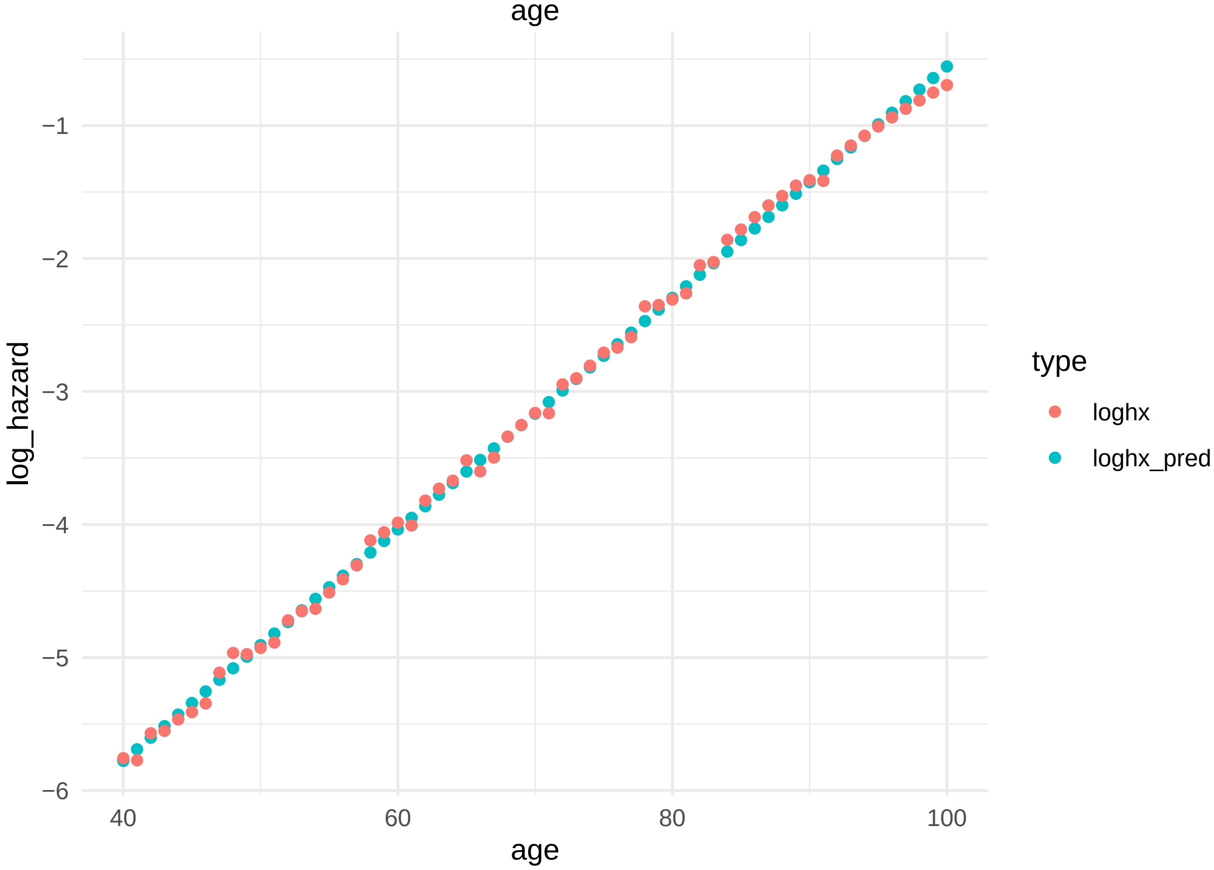
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



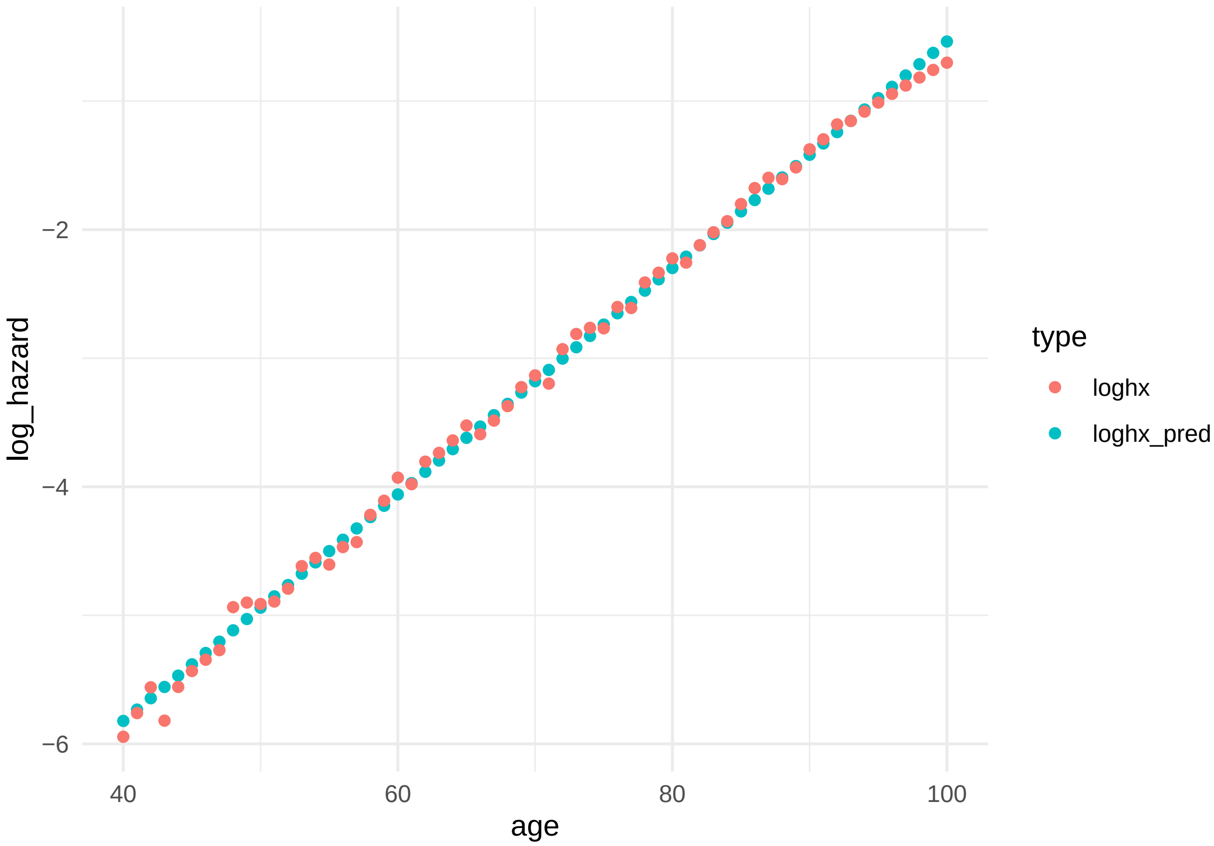
log_hazard
−1
−2
−3
−4
−5
−6
40
60
80
100
age
type
loghx
loghx_pred

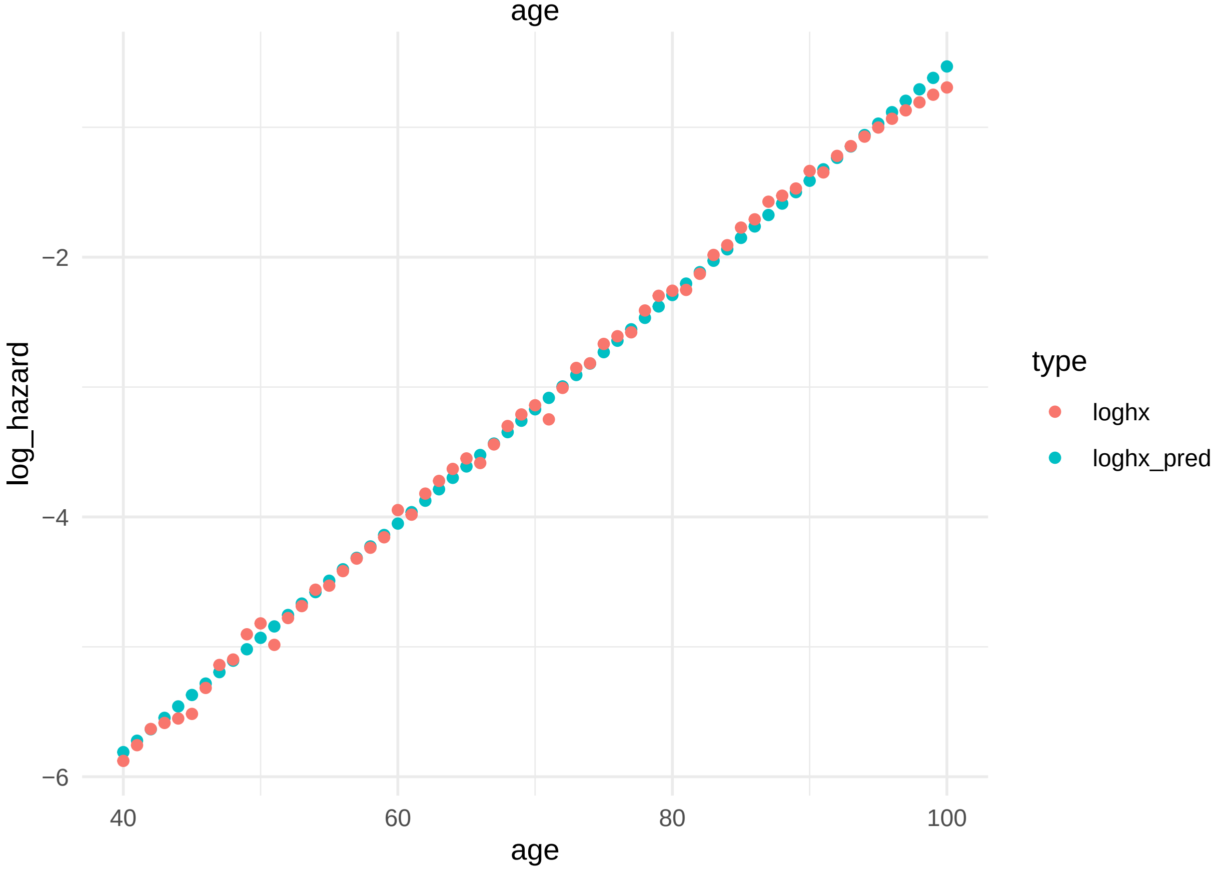


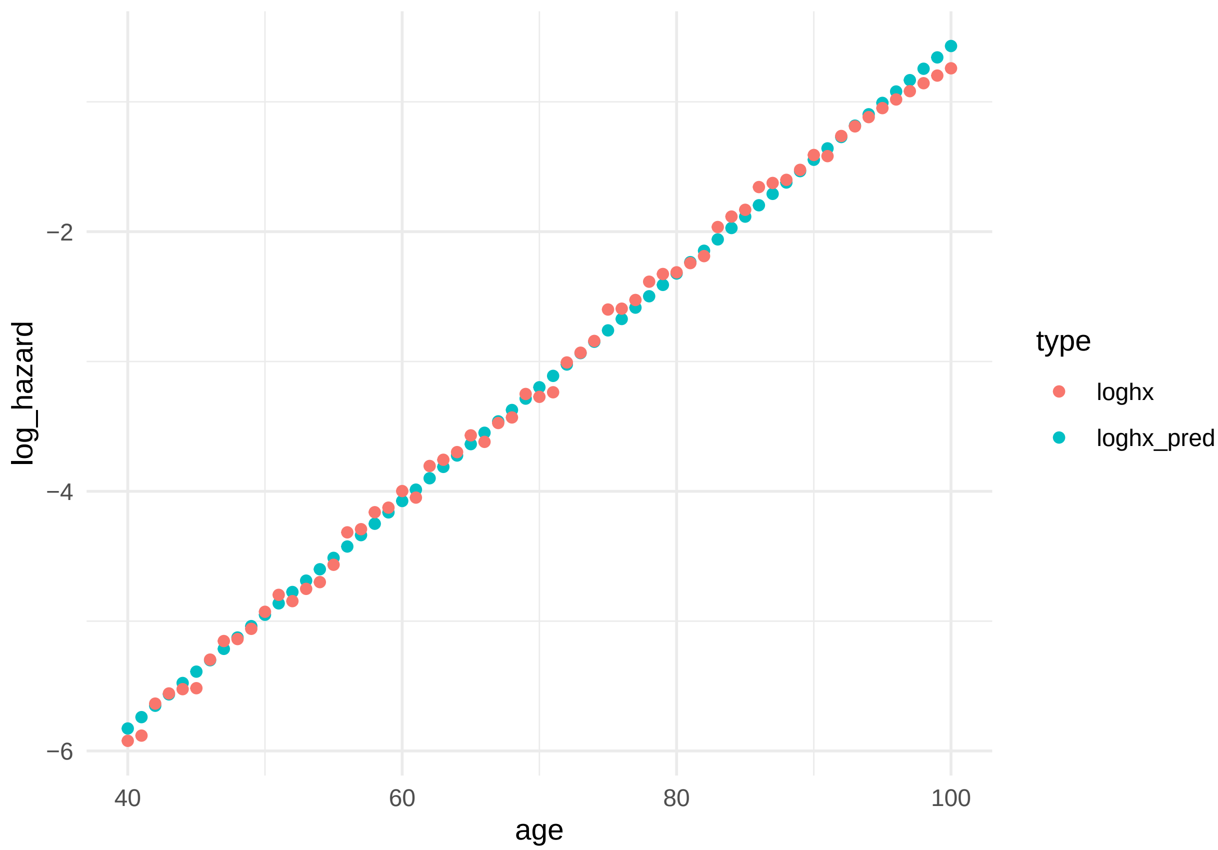
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



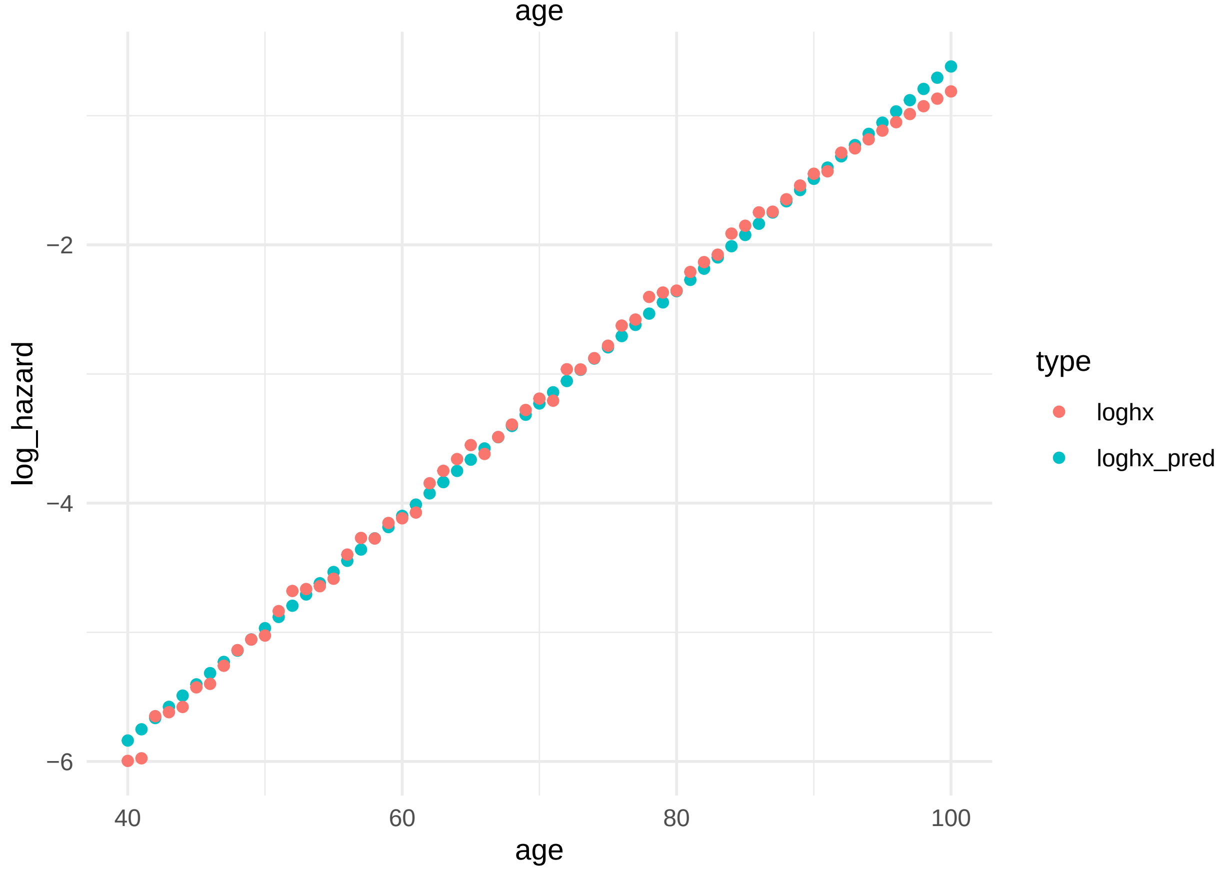
log_hazard
−1
−2
−3
−4
−5
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

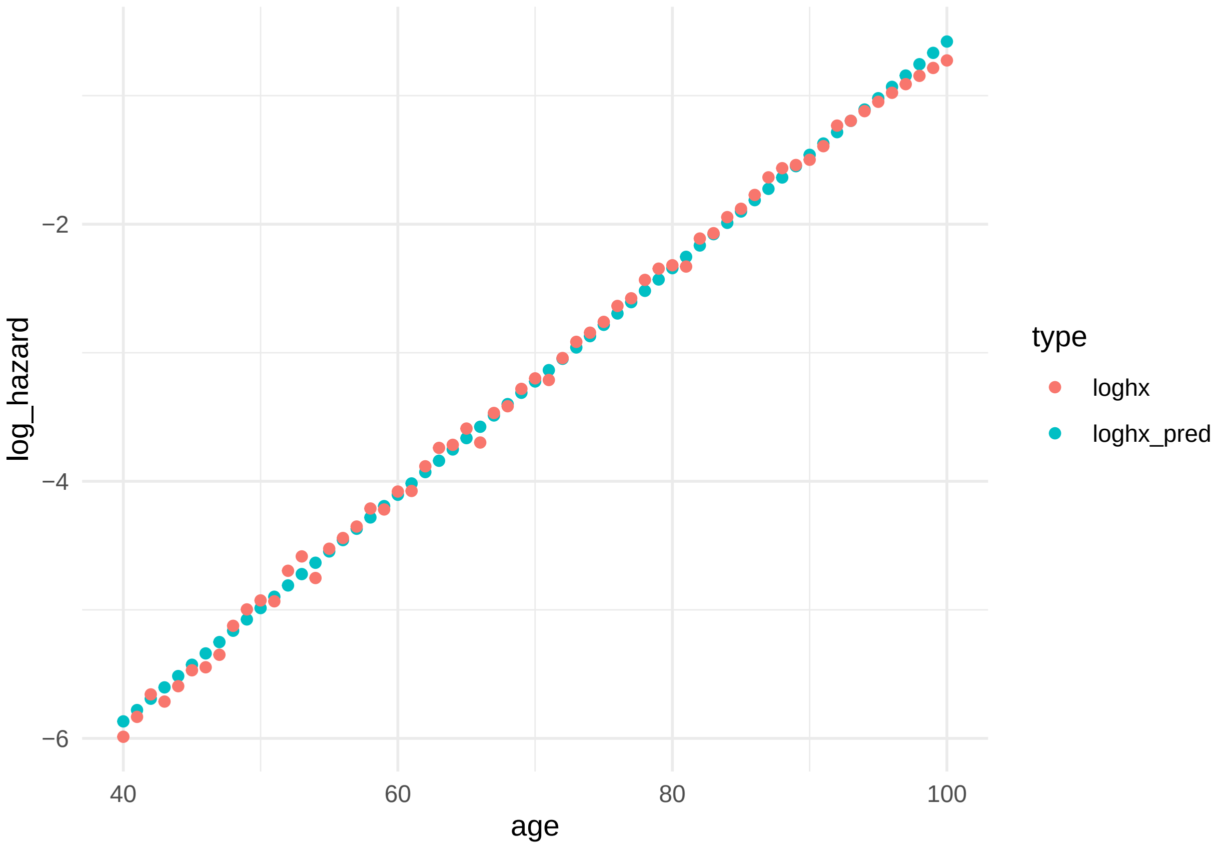

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



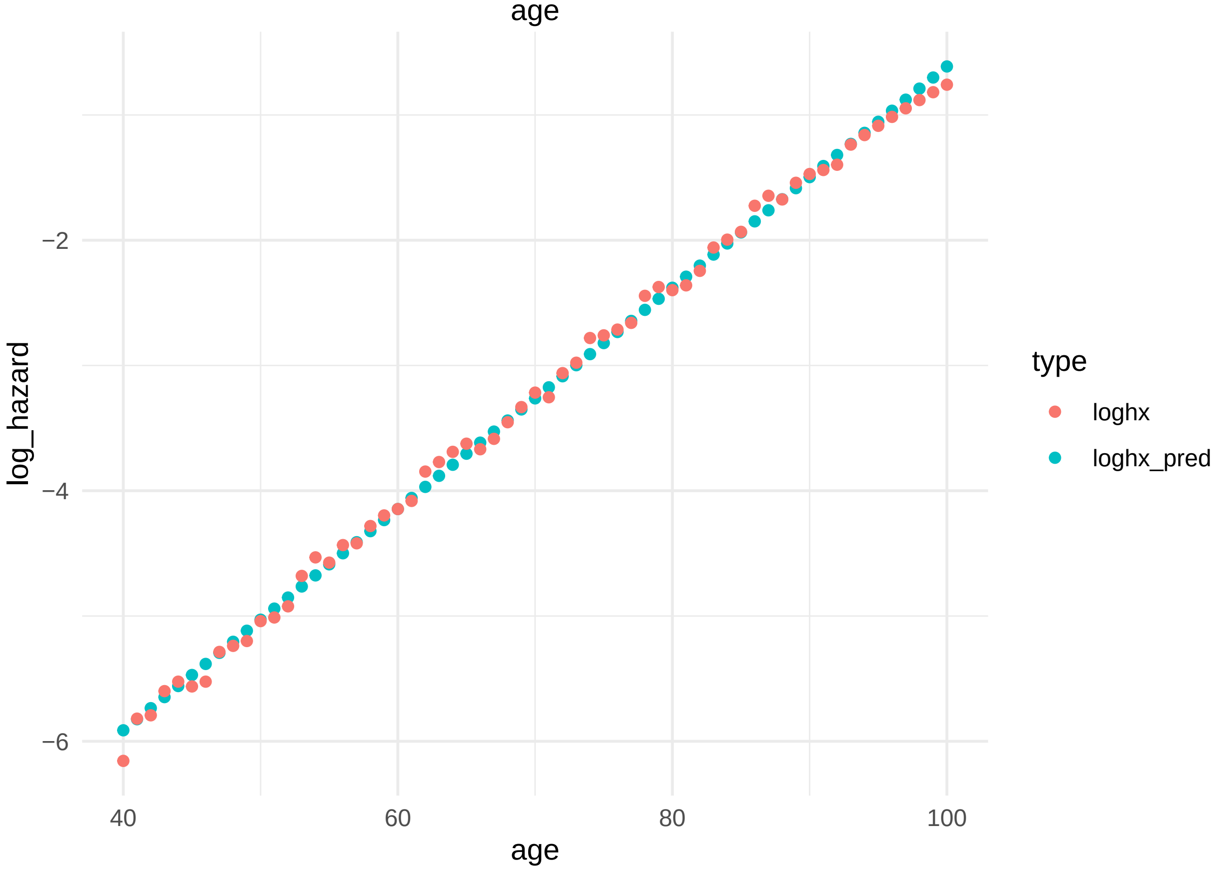
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



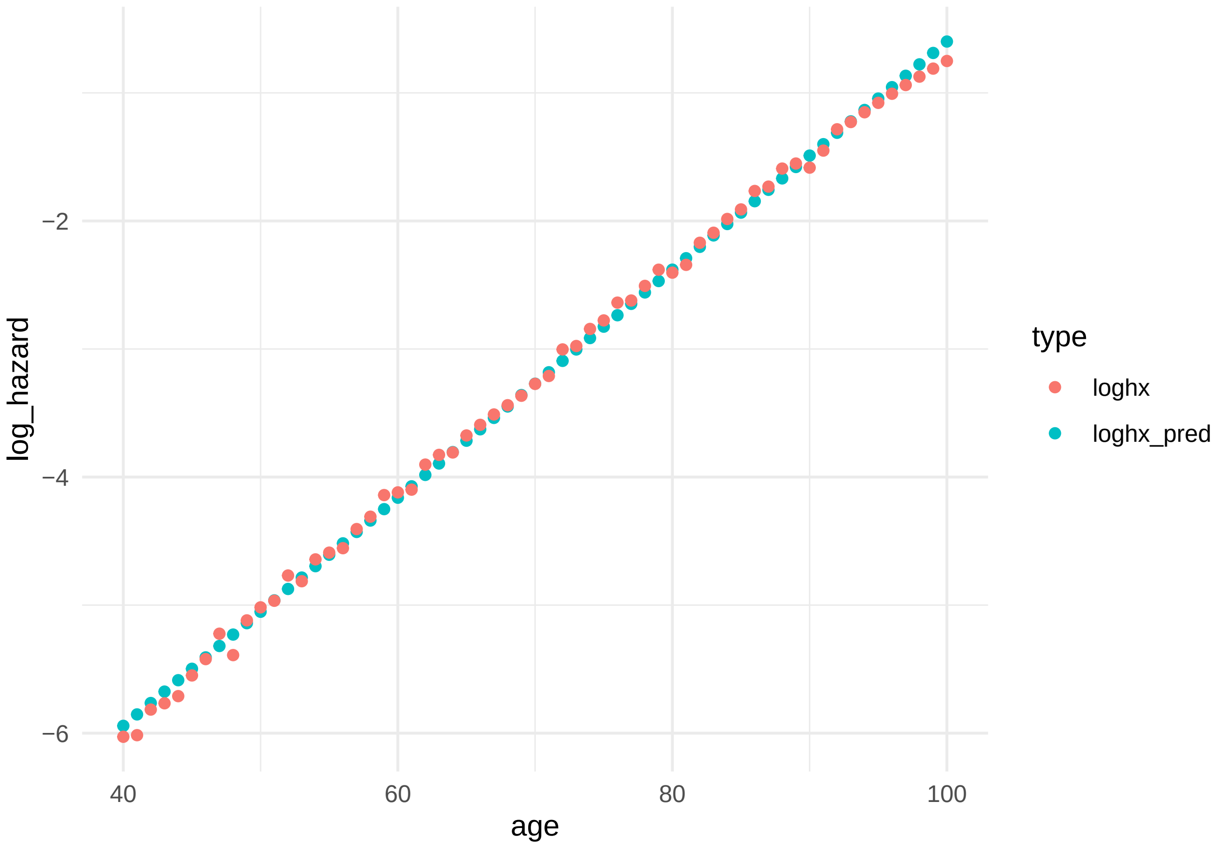
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

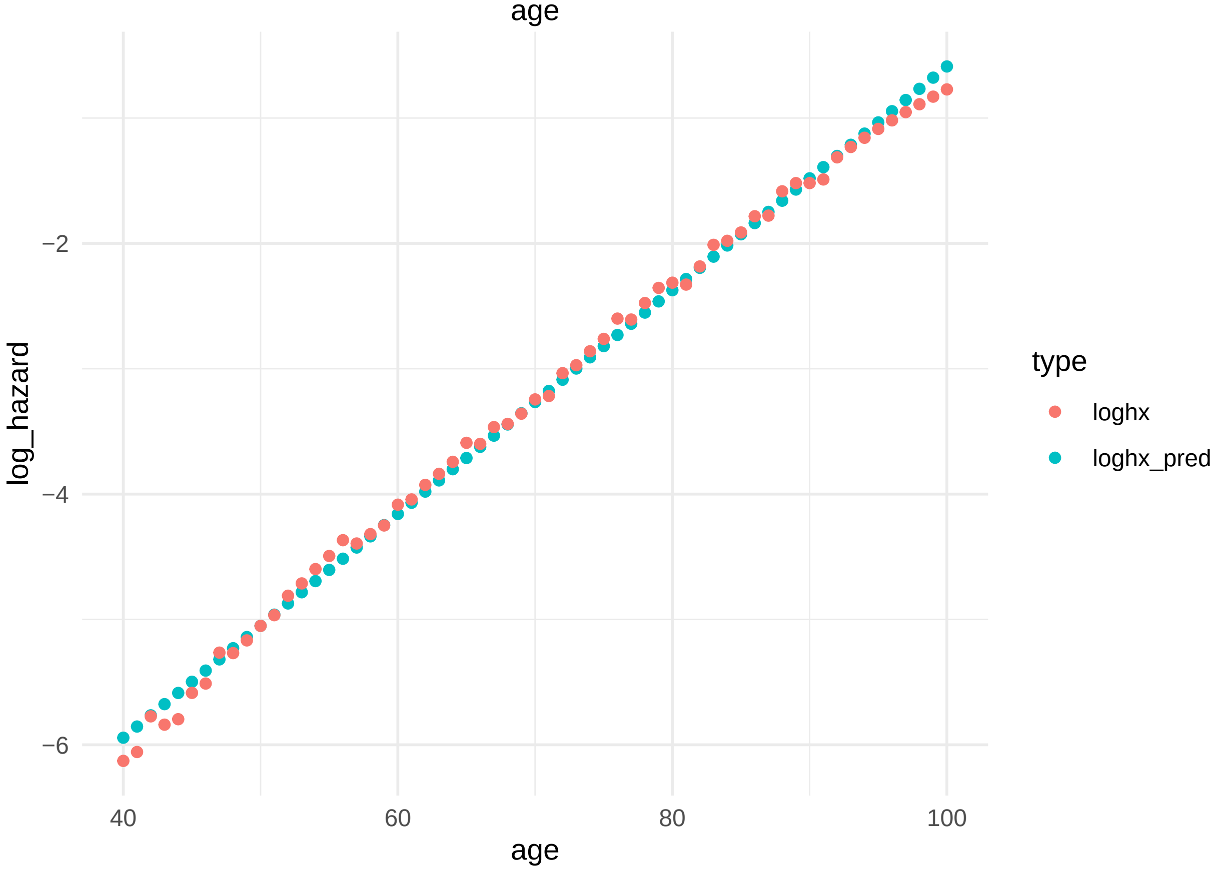


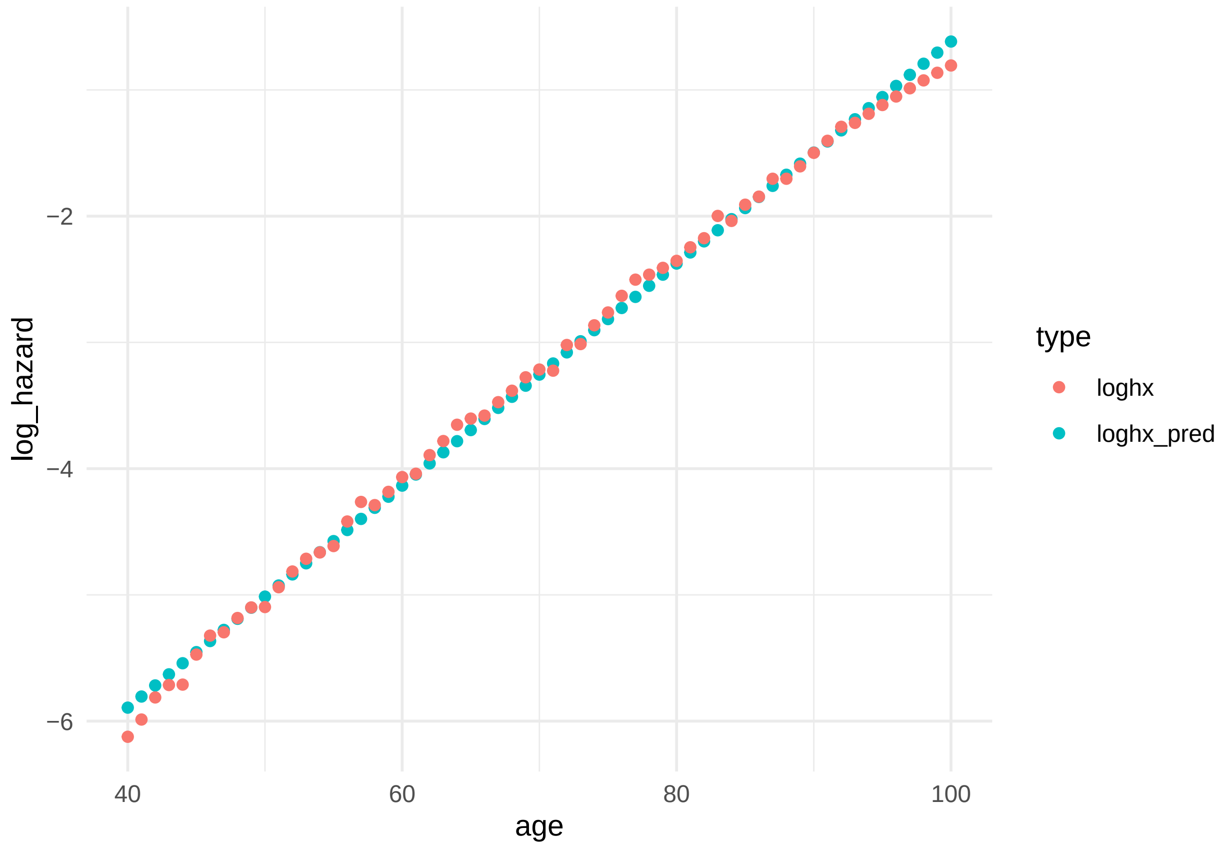
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

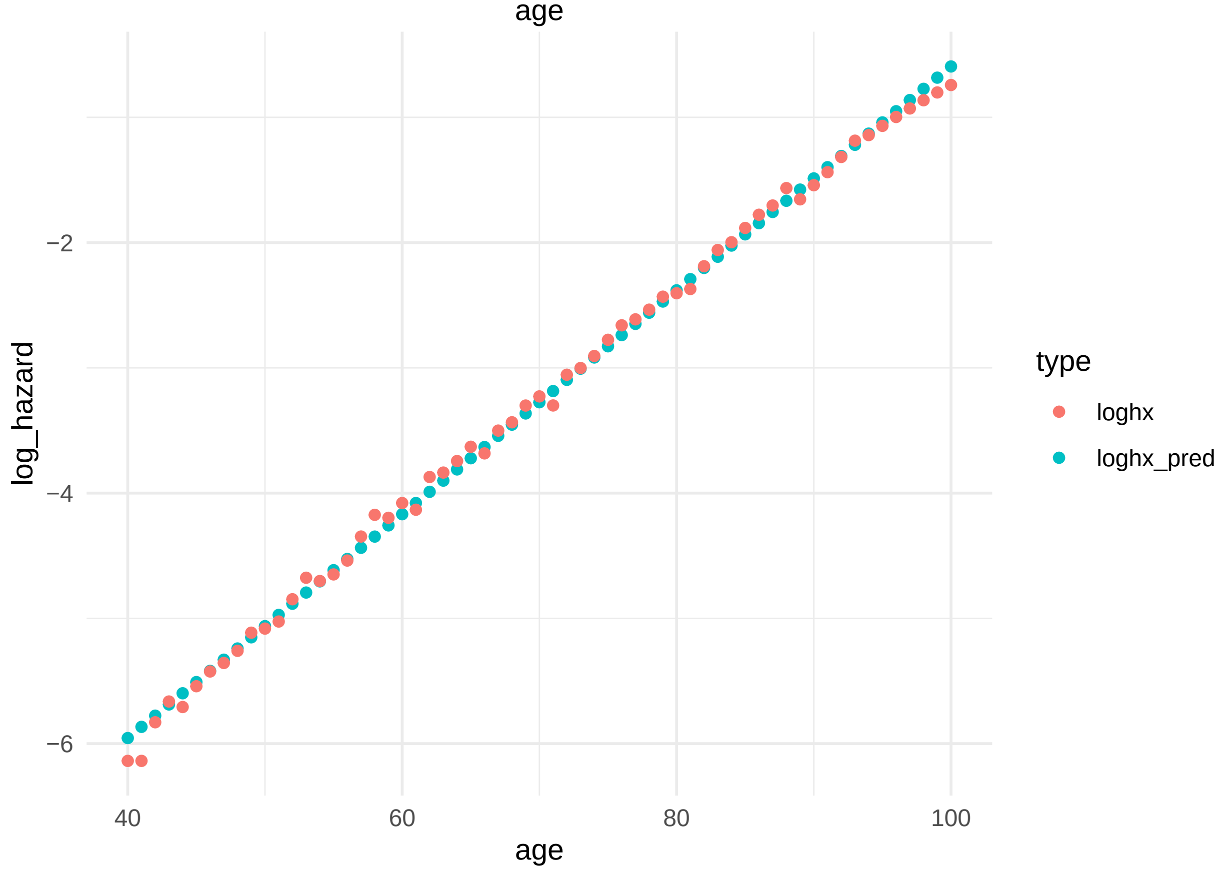


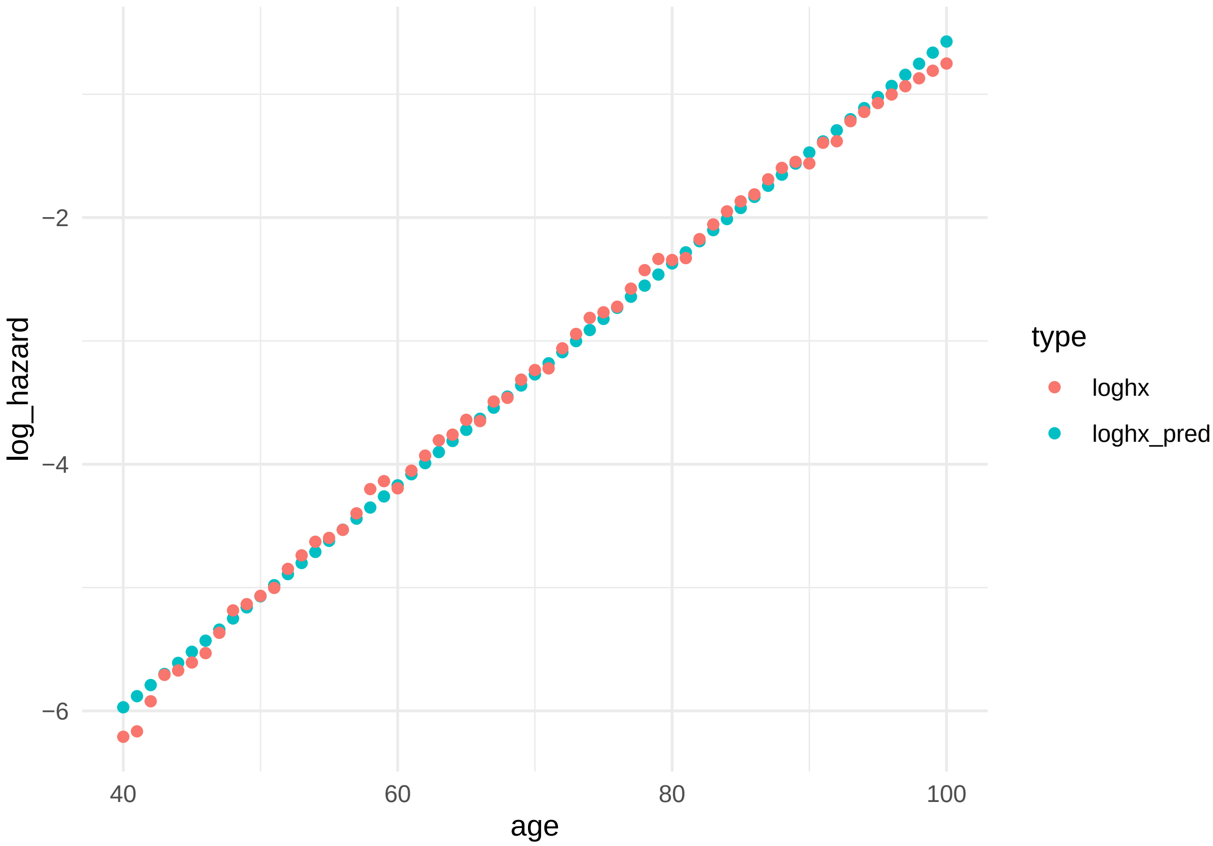
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

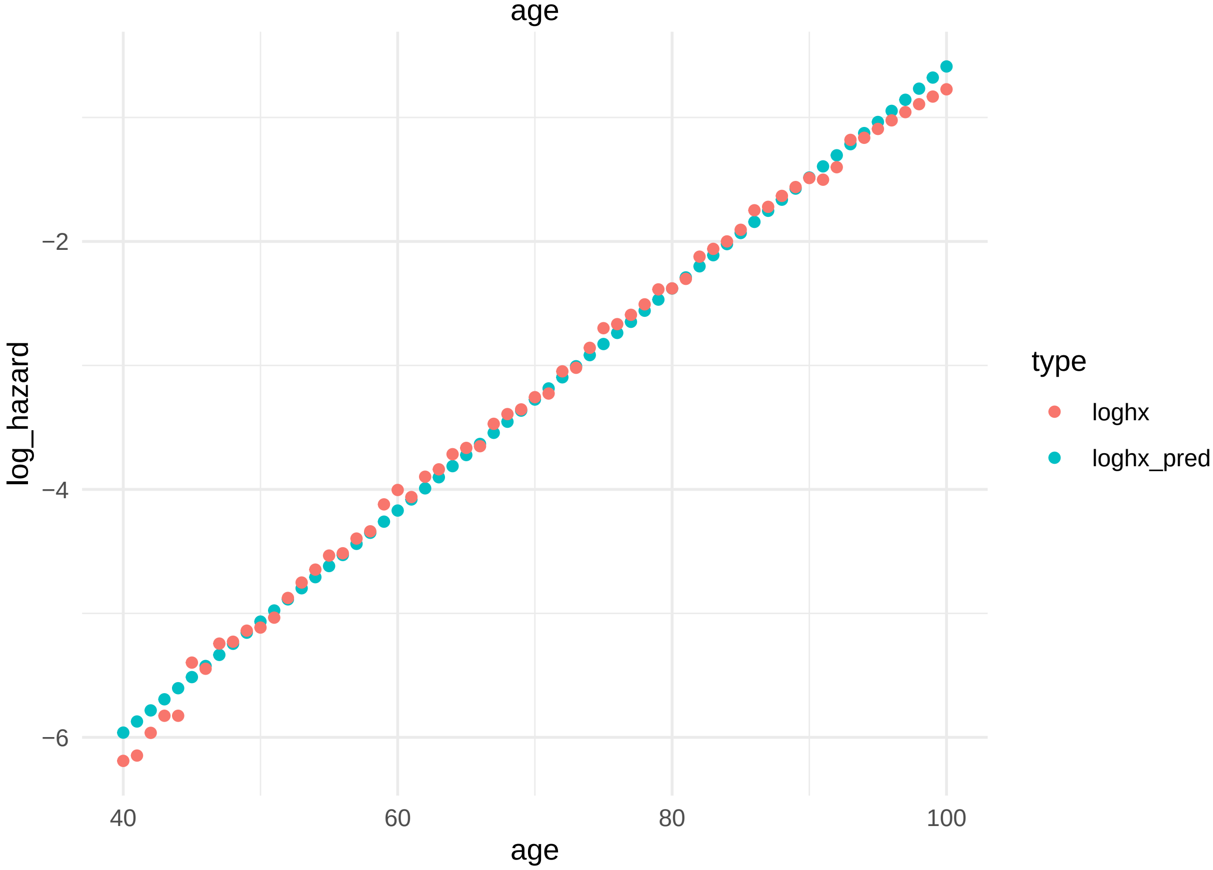

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

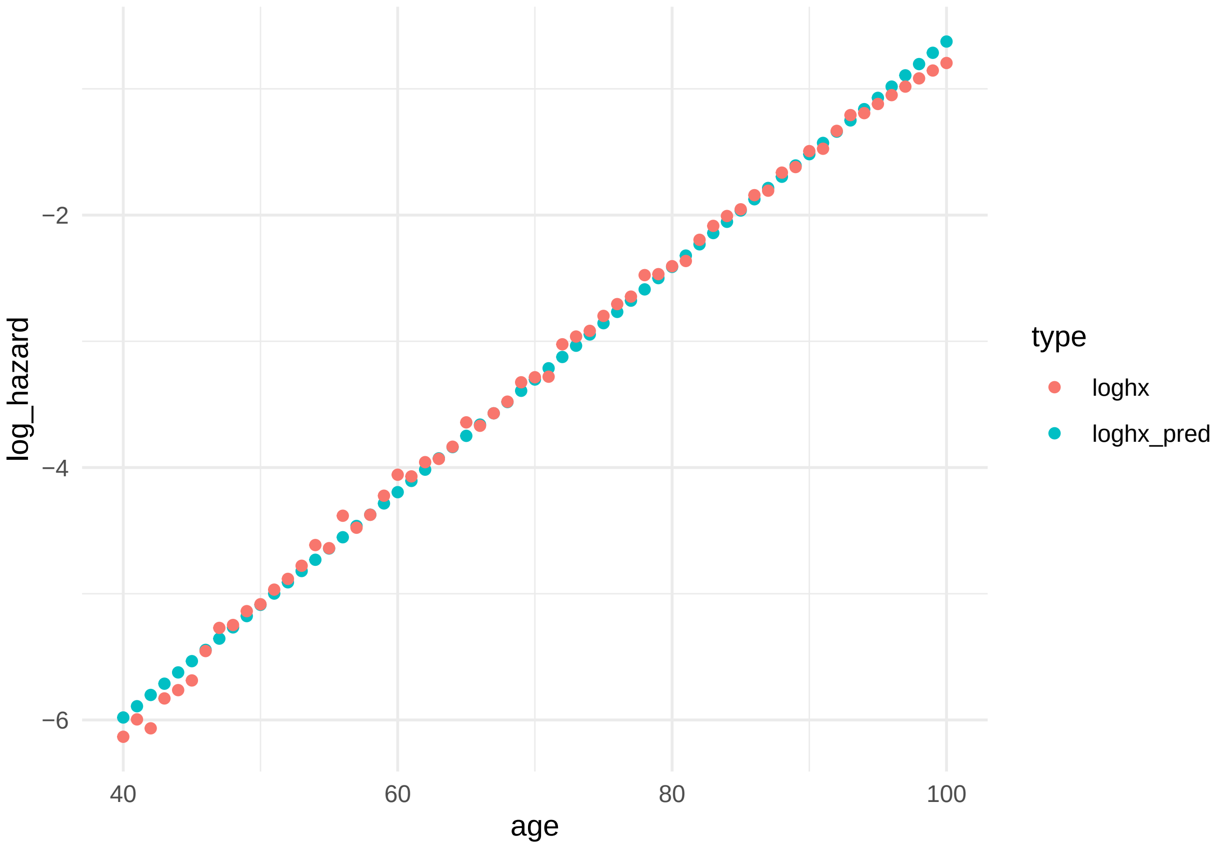

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

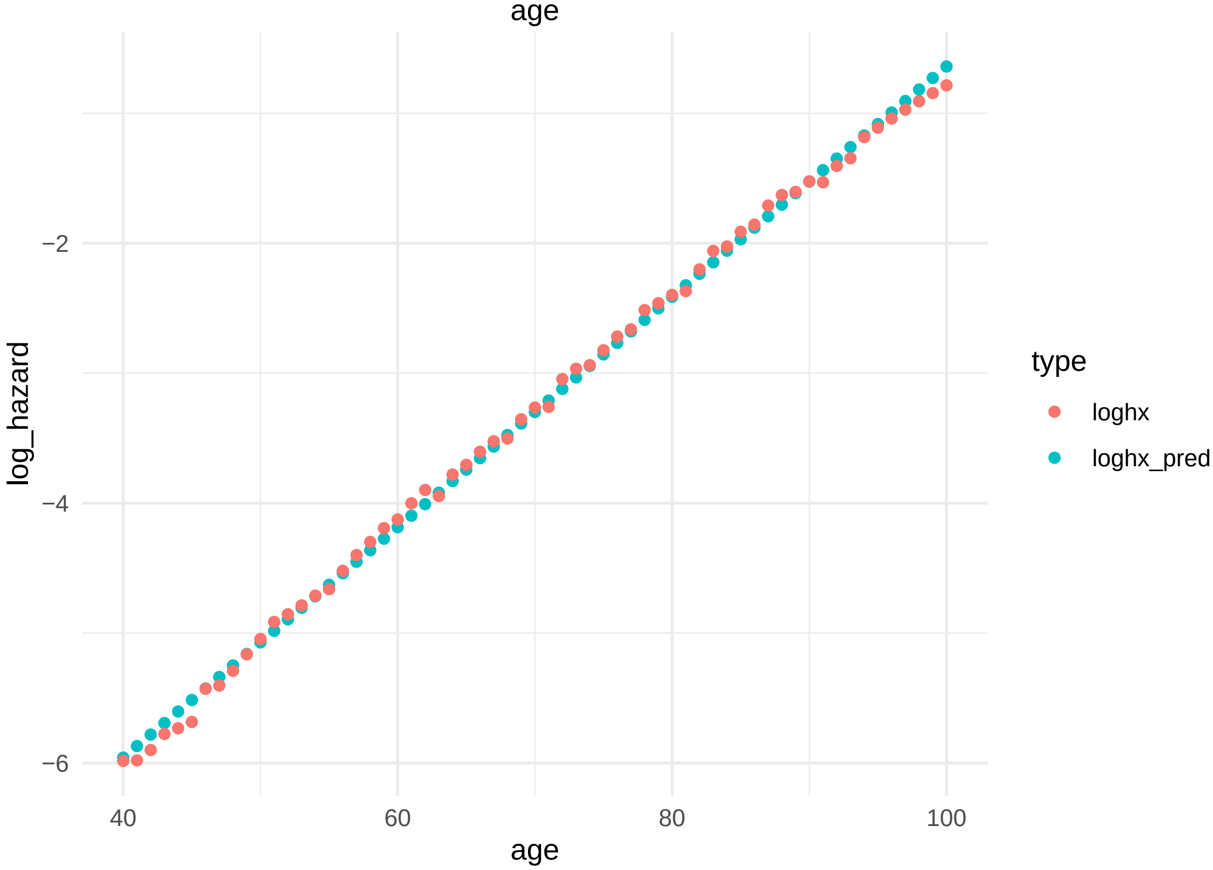

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

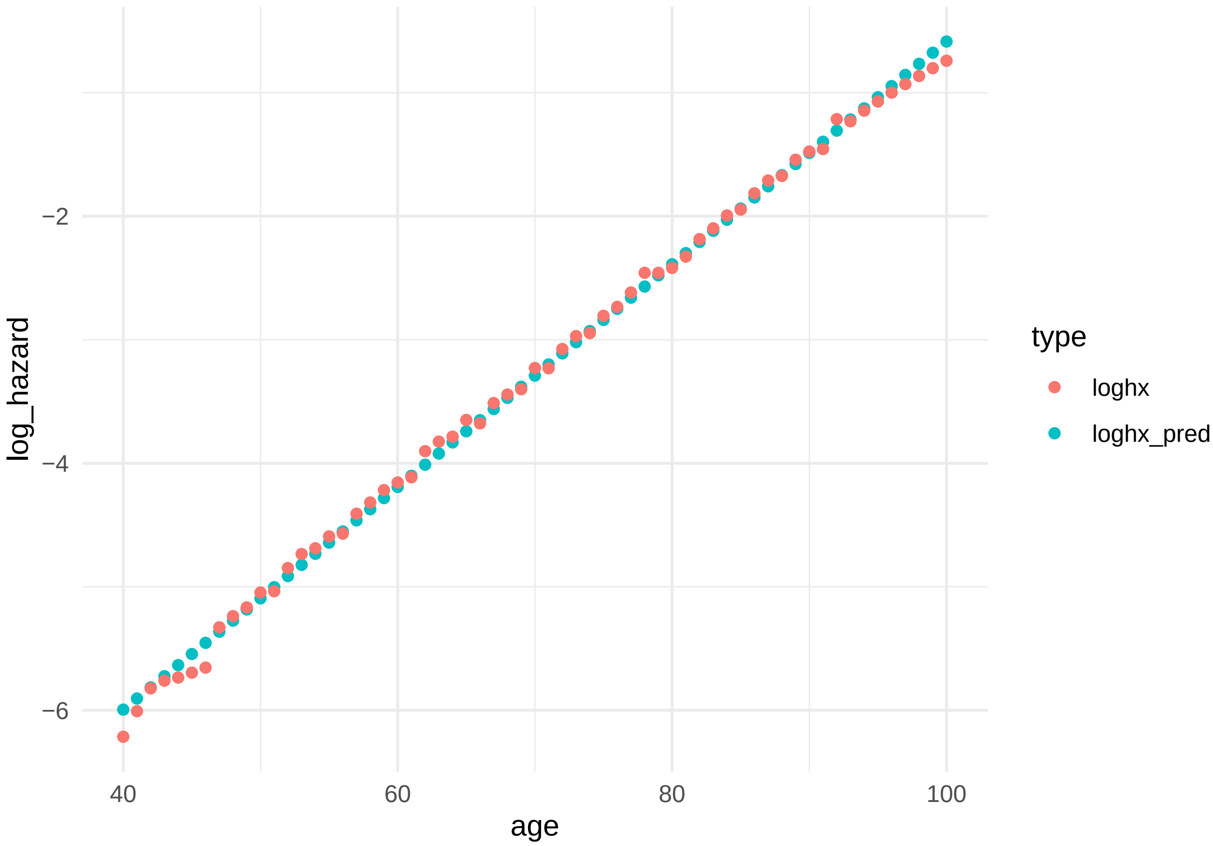

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

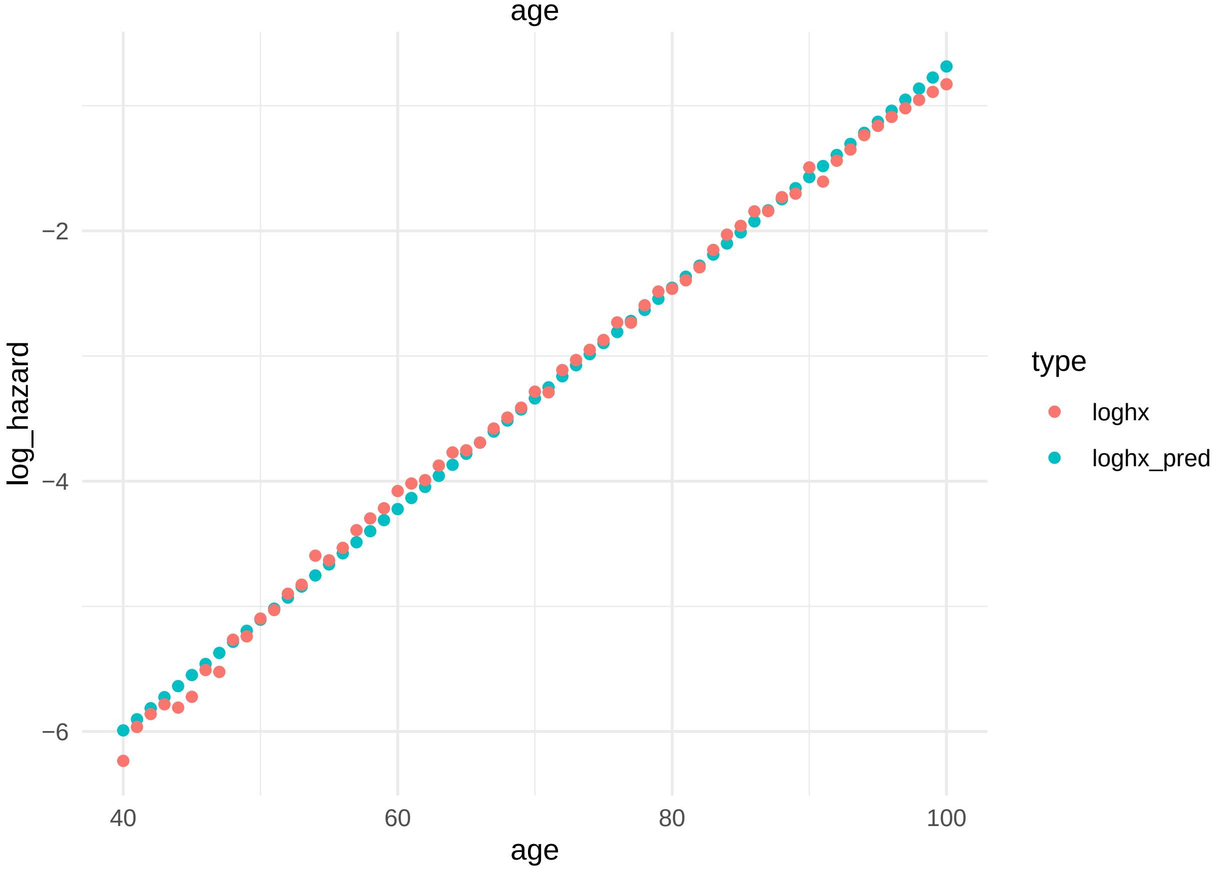

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

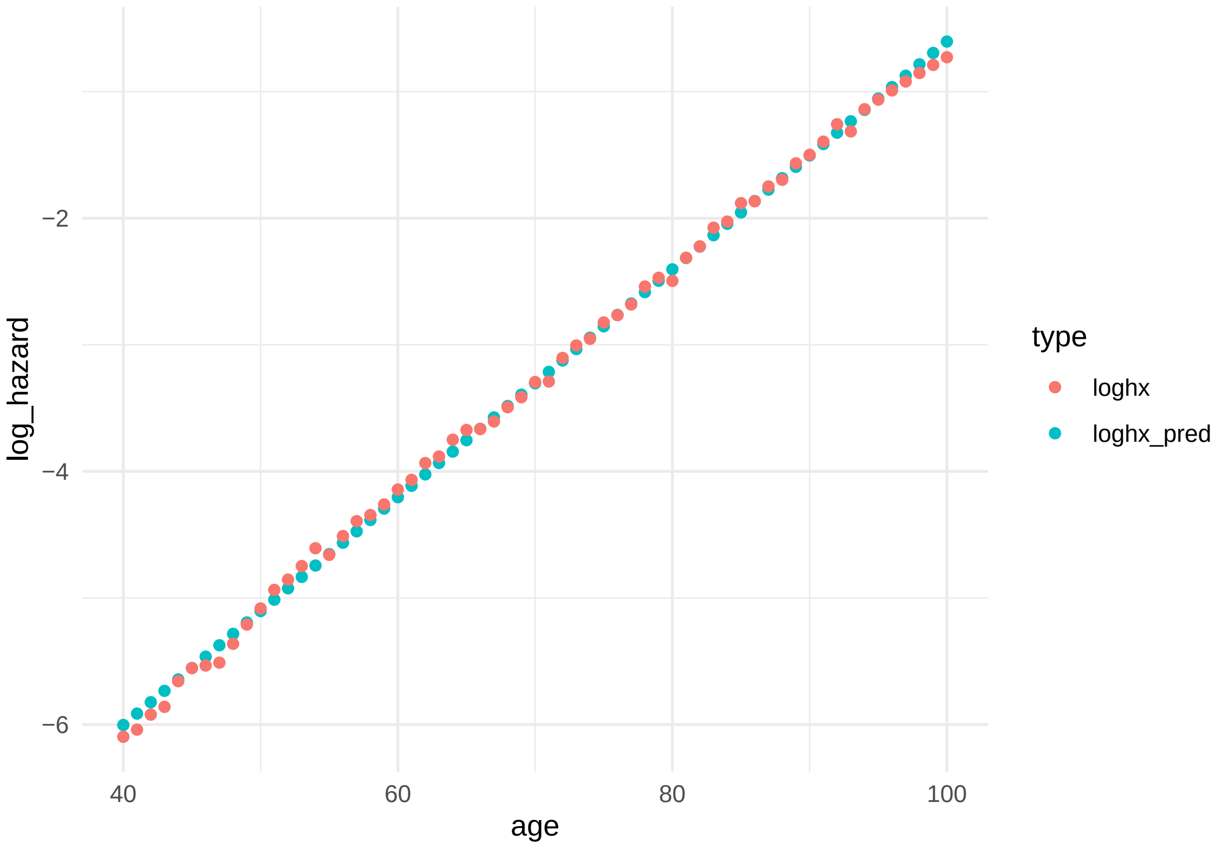

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

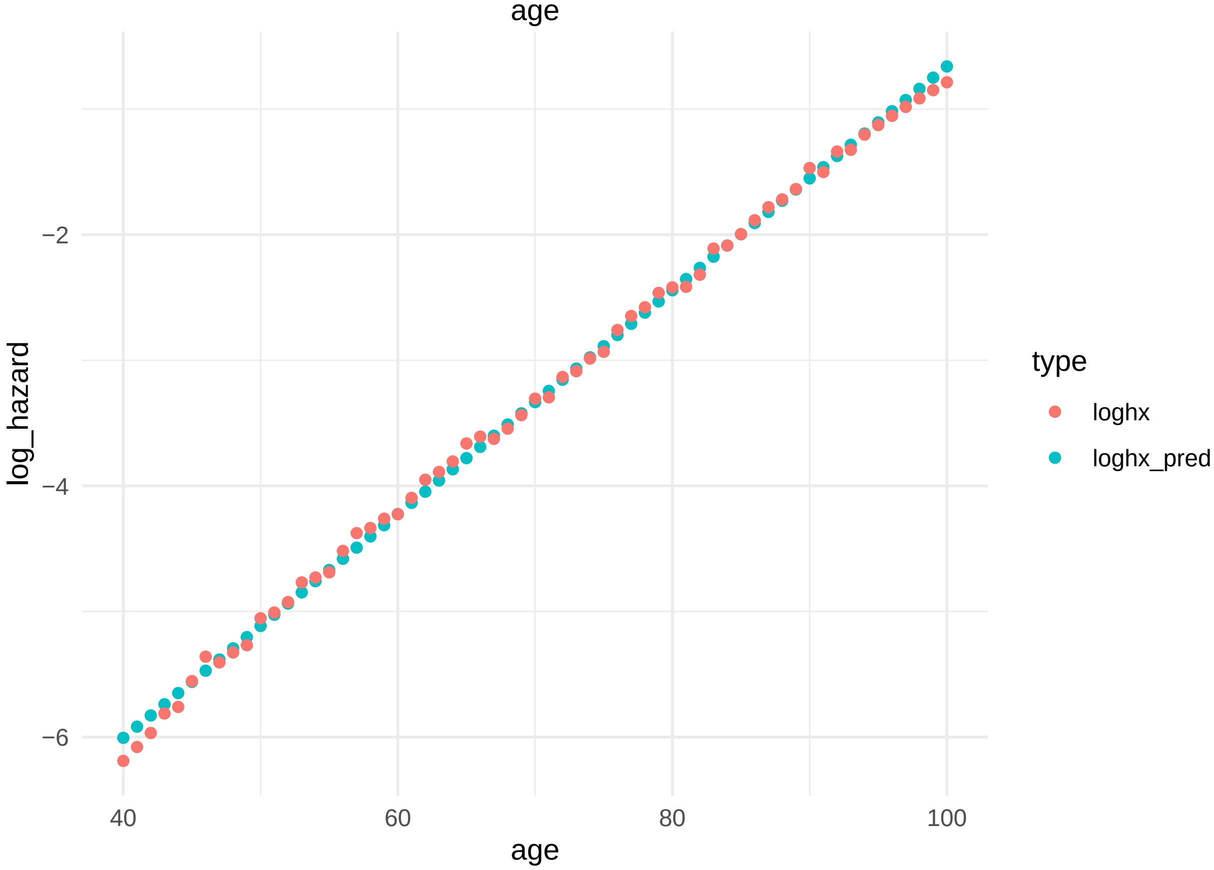

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

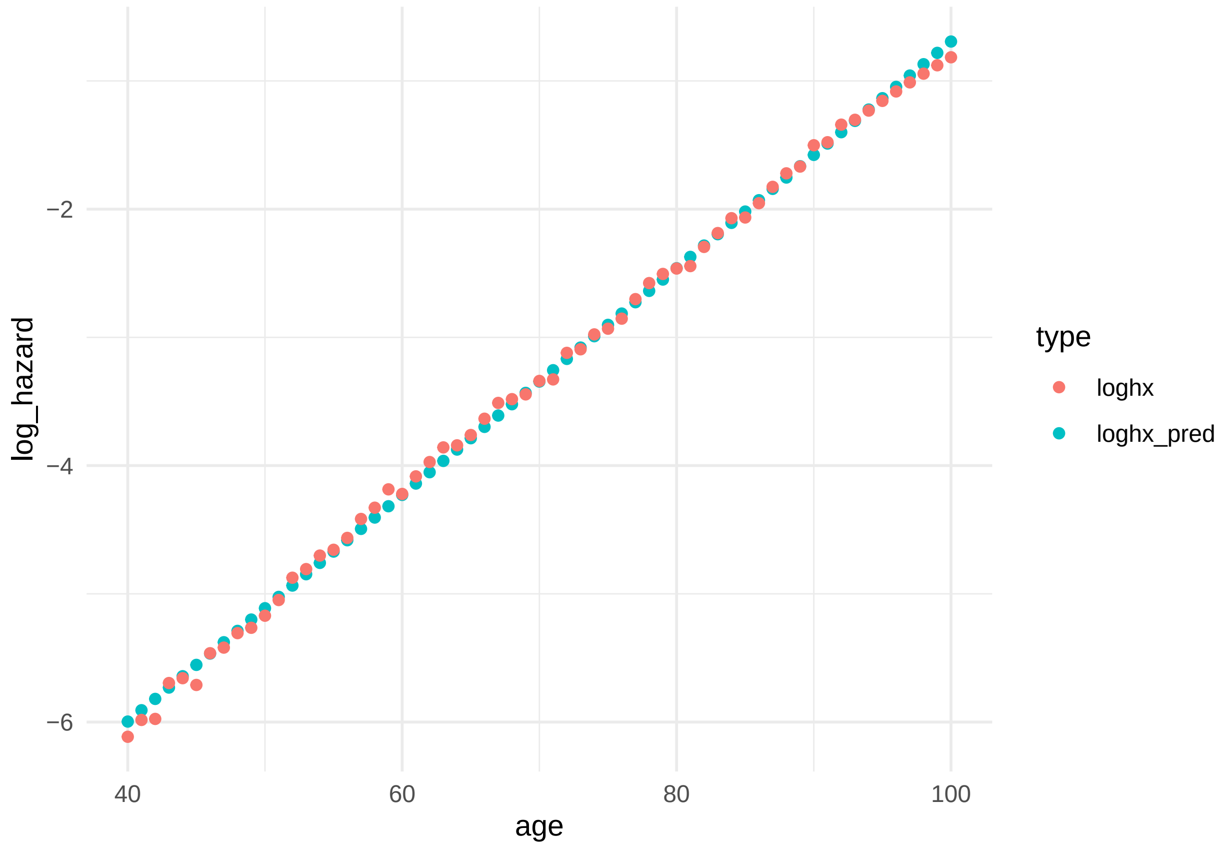

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

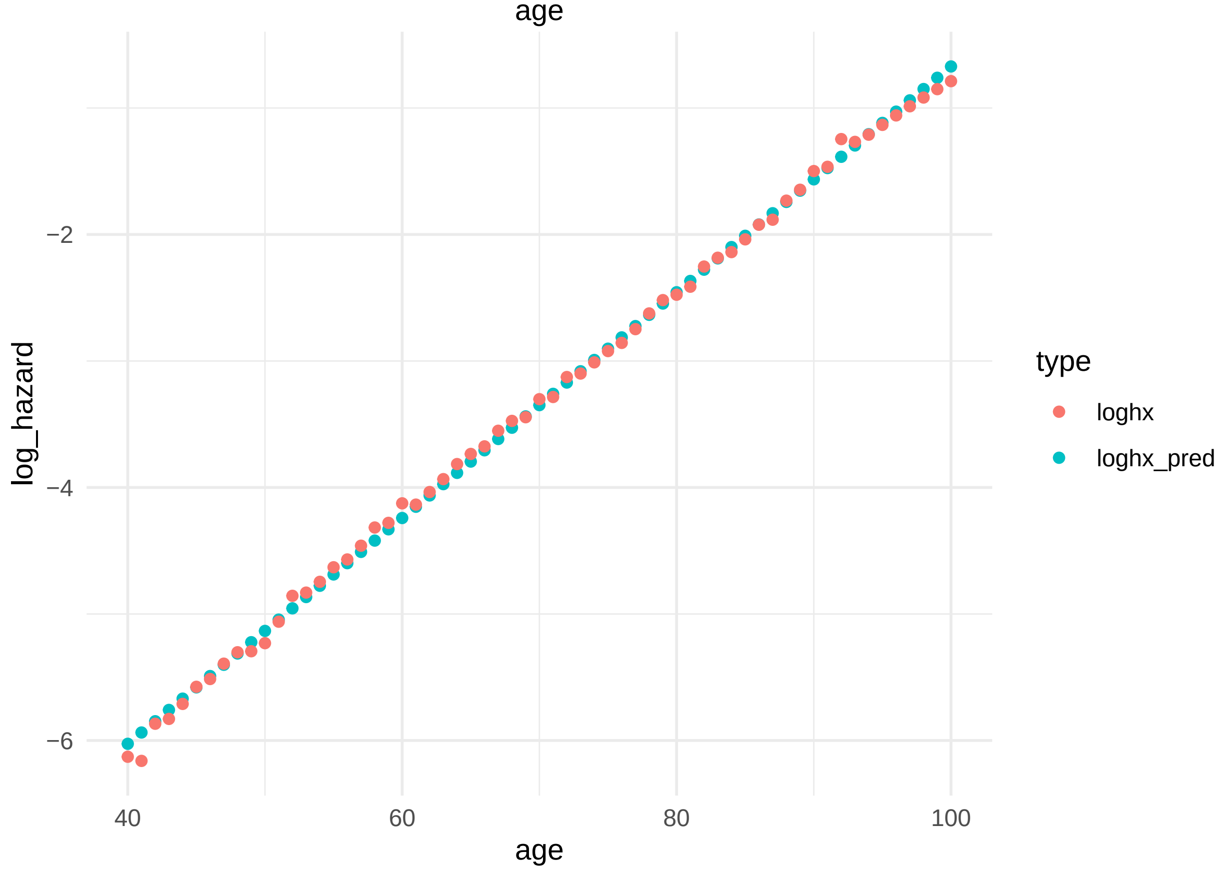

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

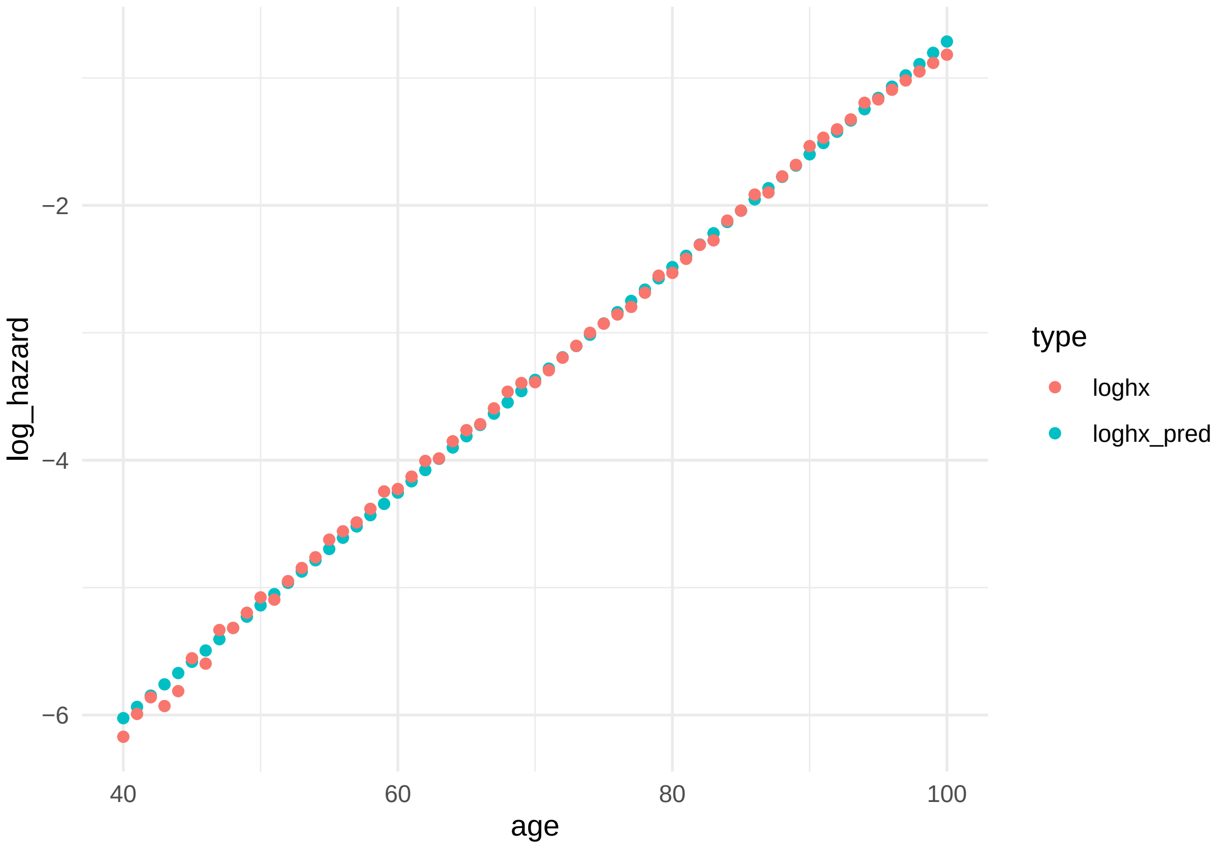

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

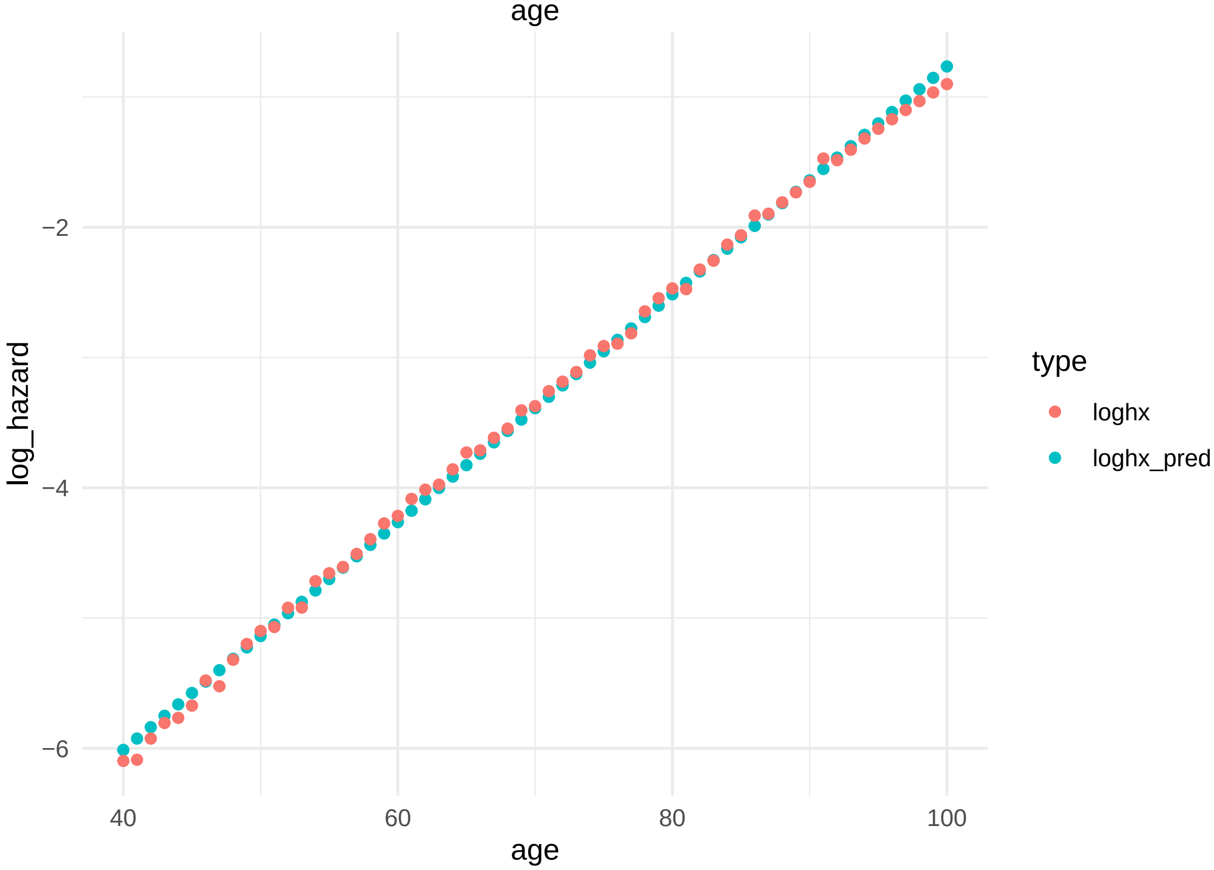

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

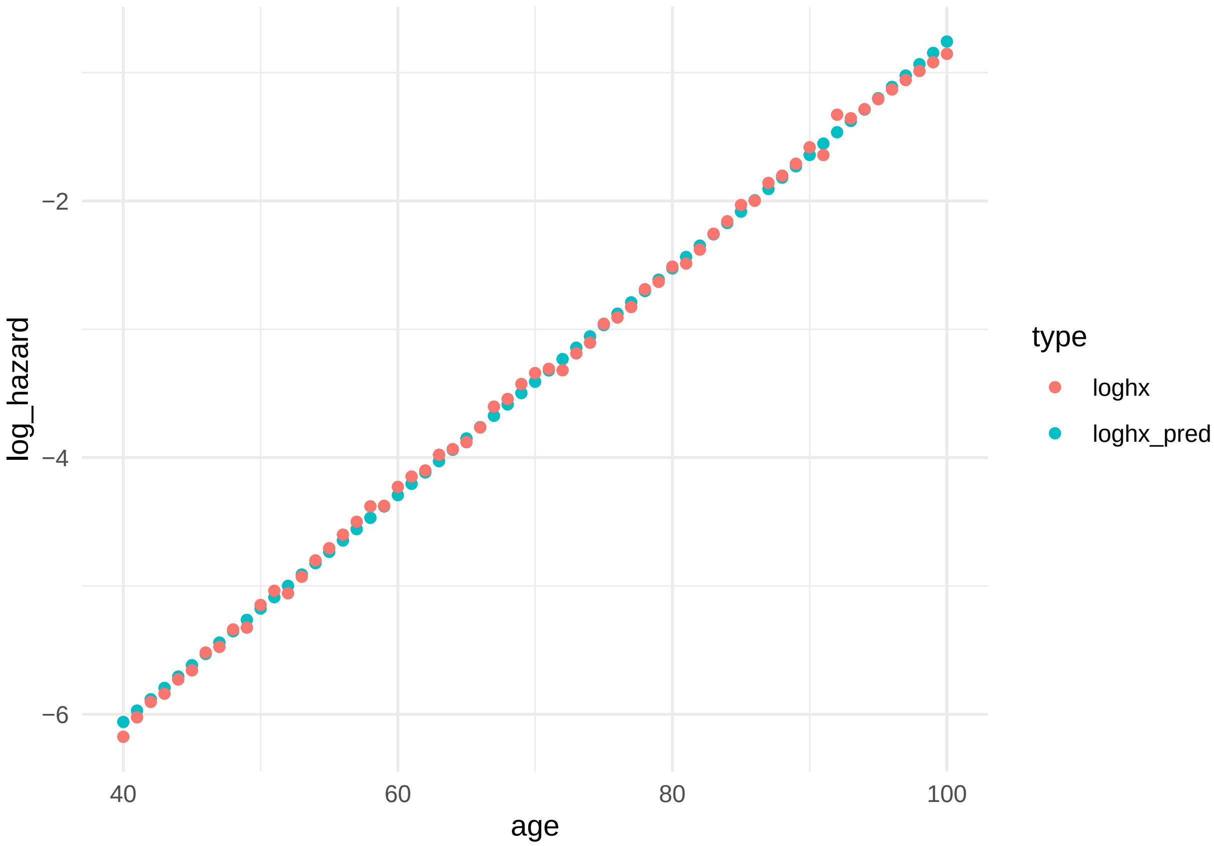

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



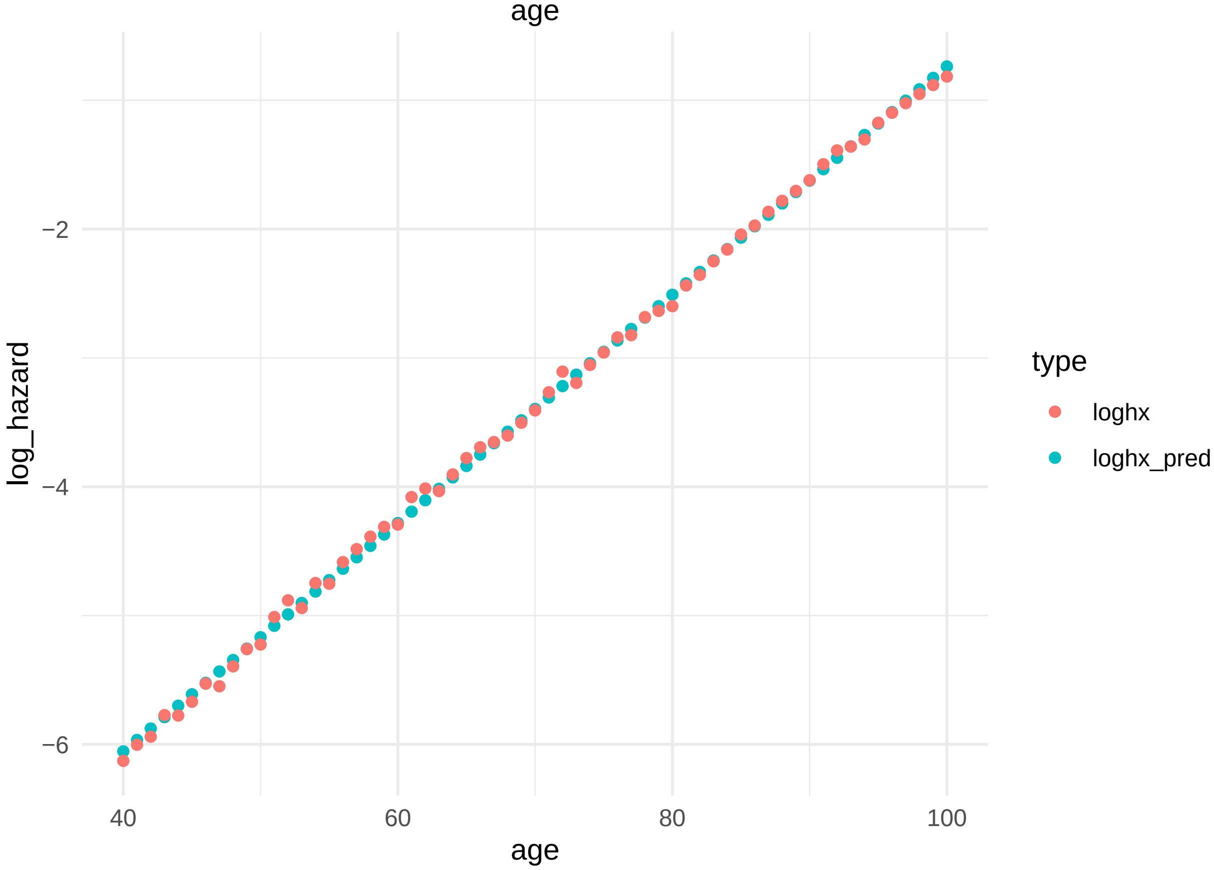
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



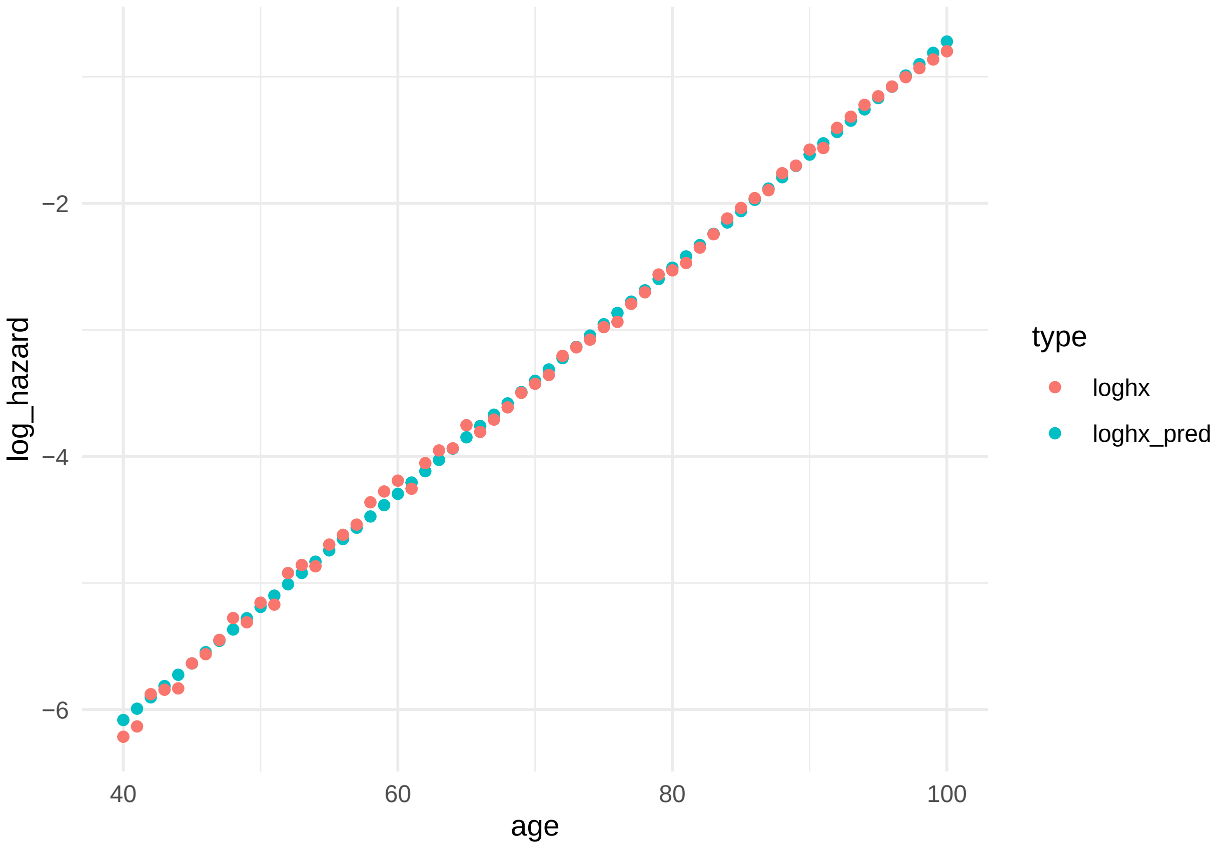
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

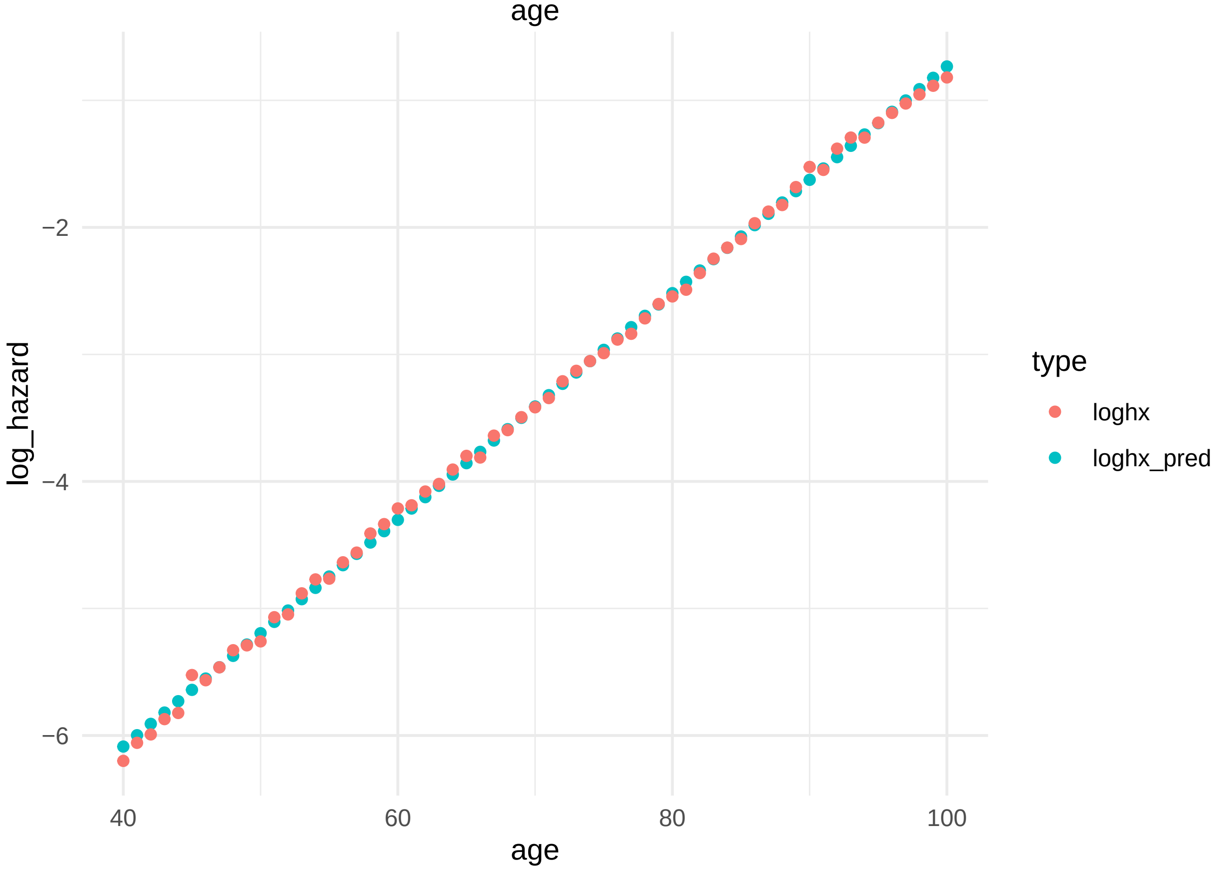


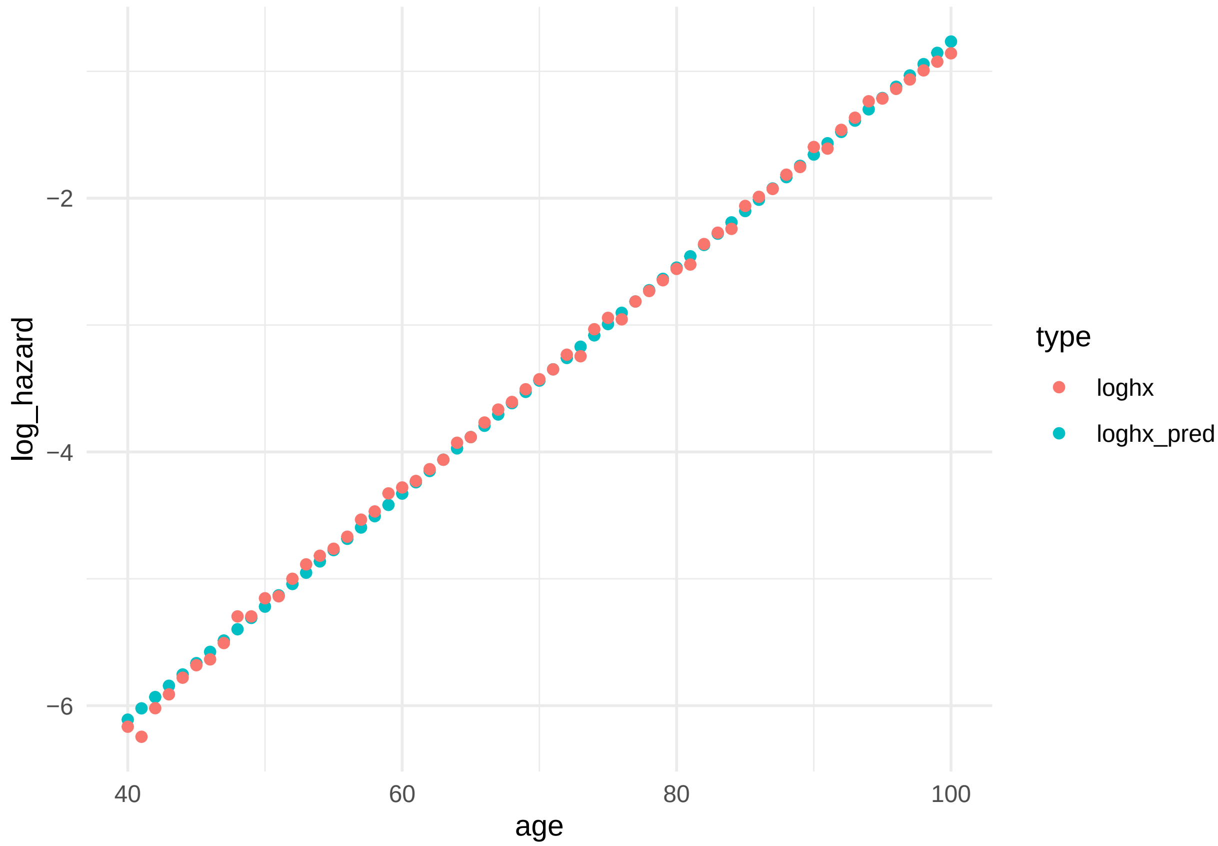
log_hazard
age
type
loghx
loghx_pred
−2
−4
−6
40
60
80
100

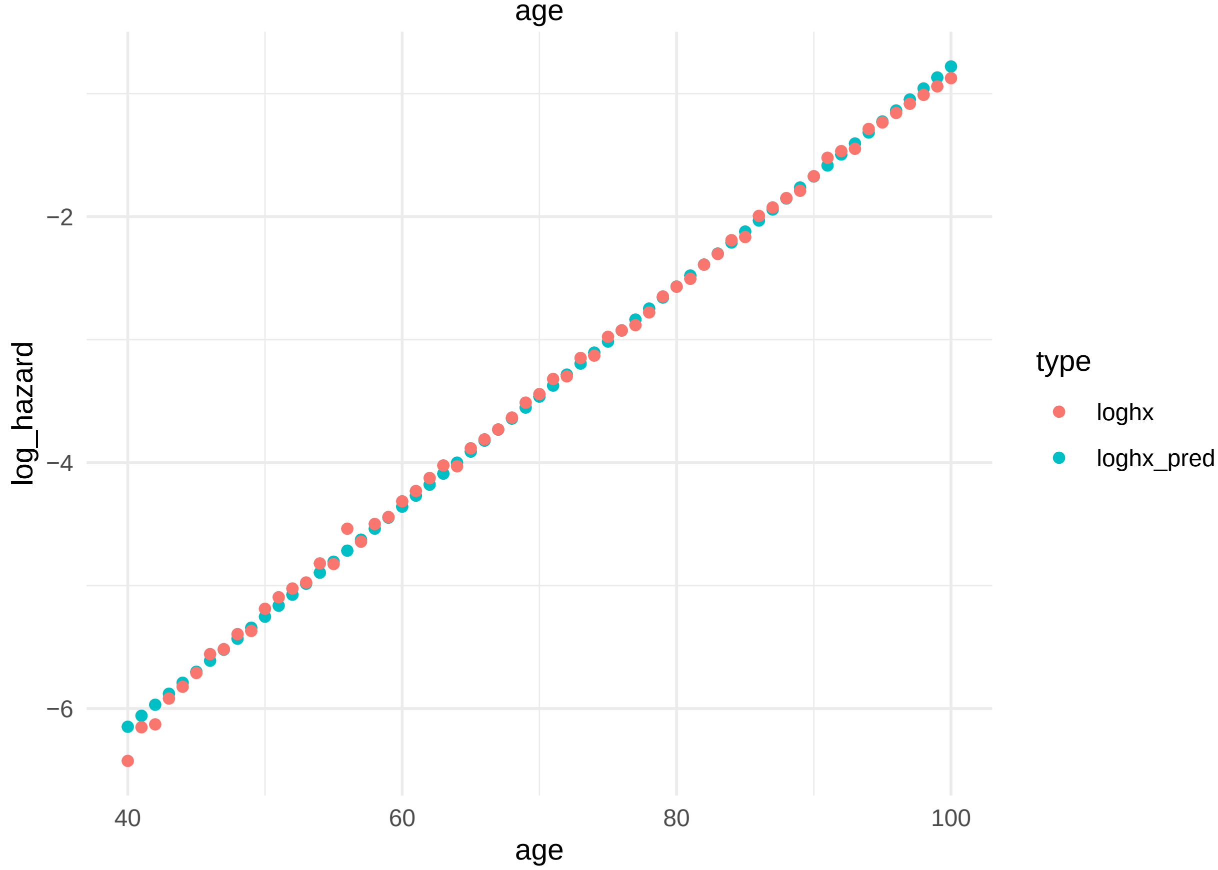


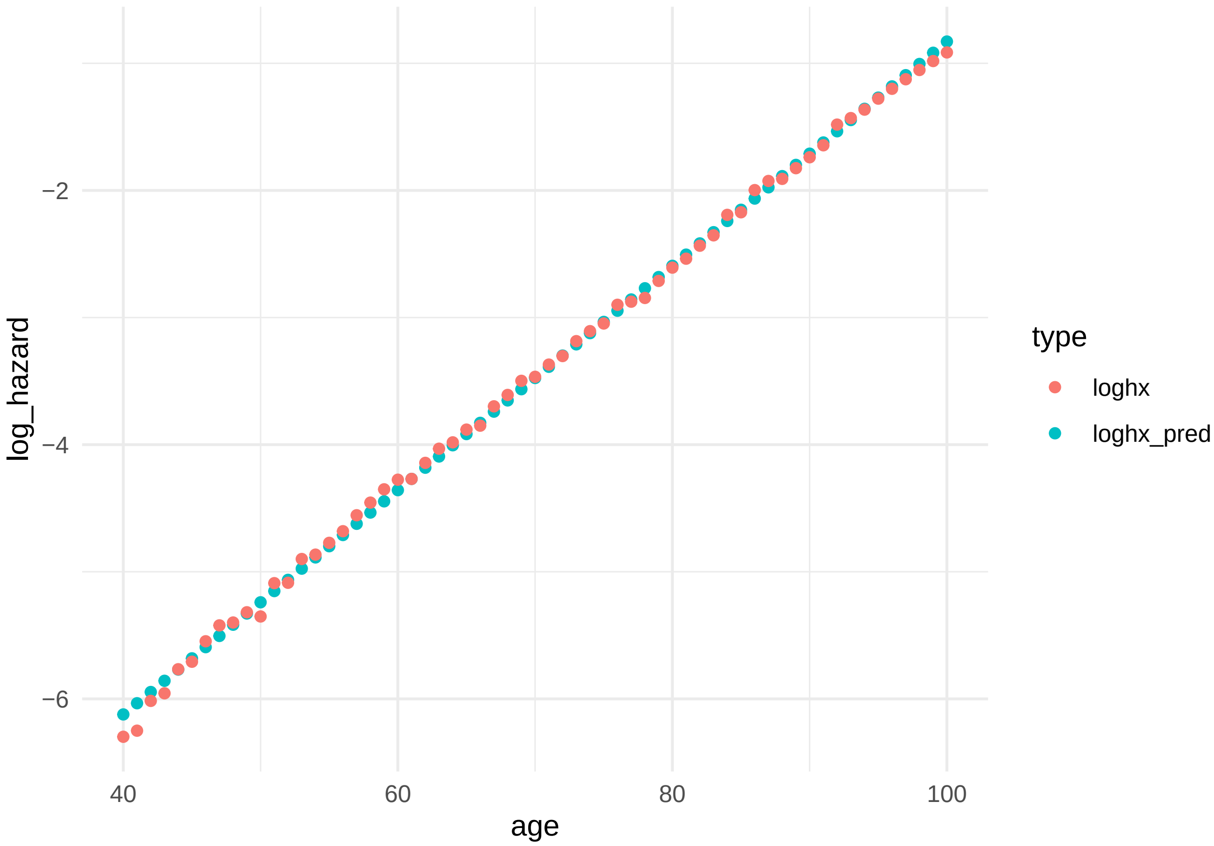
log_hazard
age
type
loghx
loghx_pred
−2
−4
−6
40
60
80
100


log_hazard
−2
−4
−6
age
40
60
80
100
type
loghx
loghx_pred

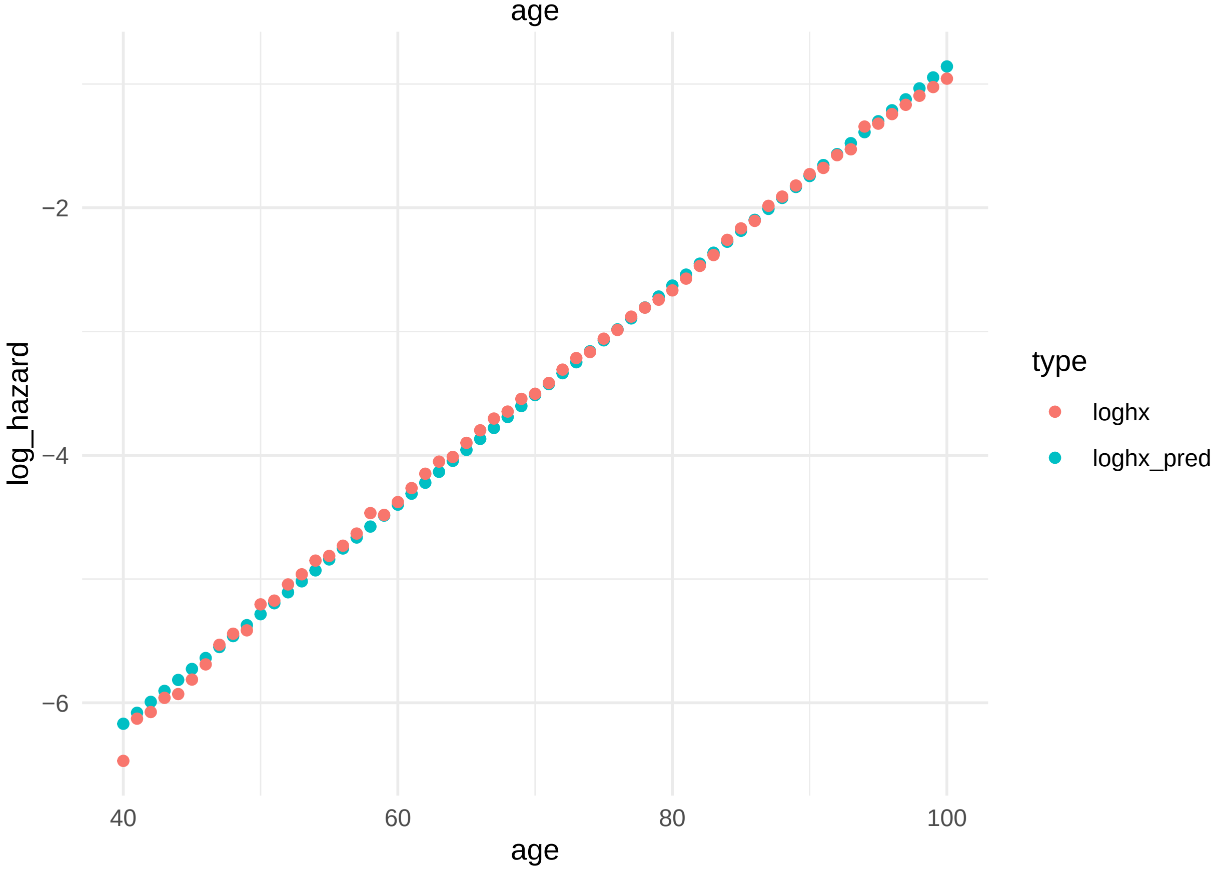

log_hazard
−2
−4
−6
age
40
60
80
100
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

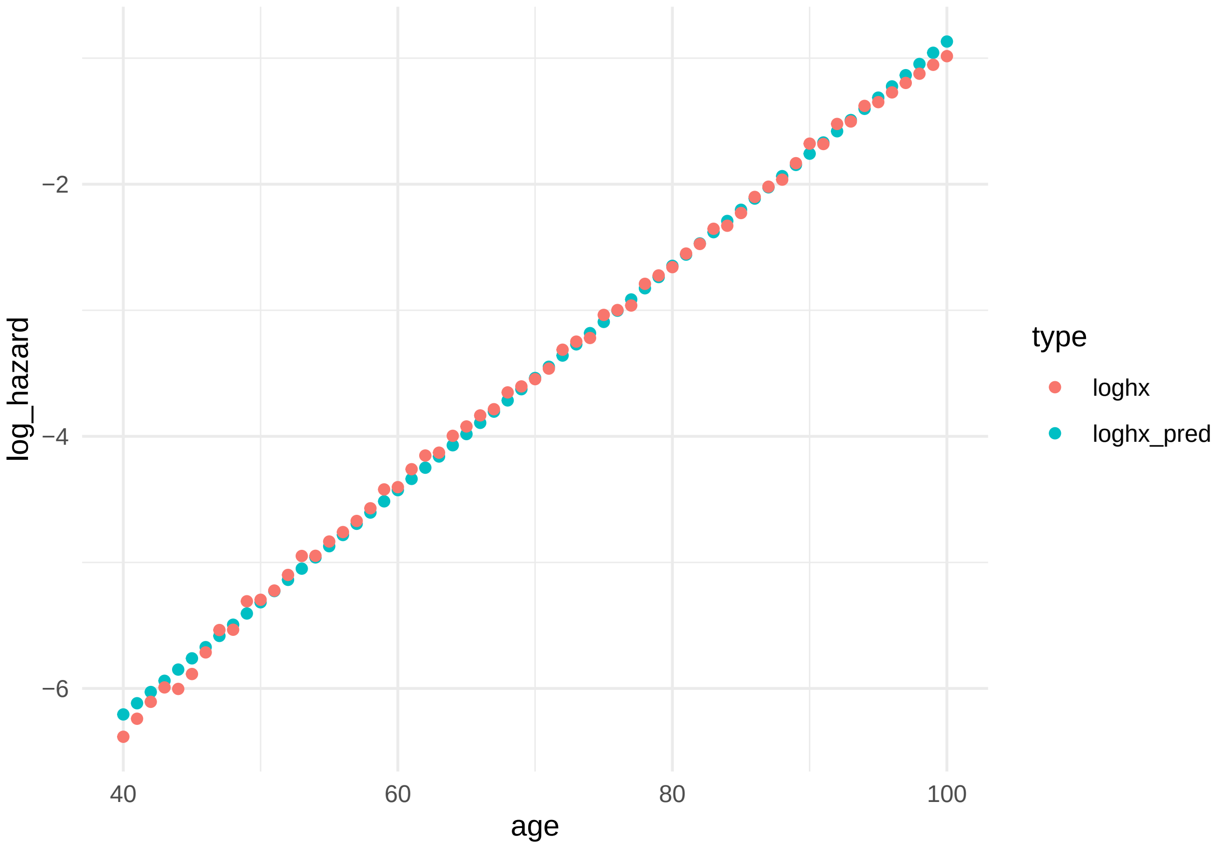

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

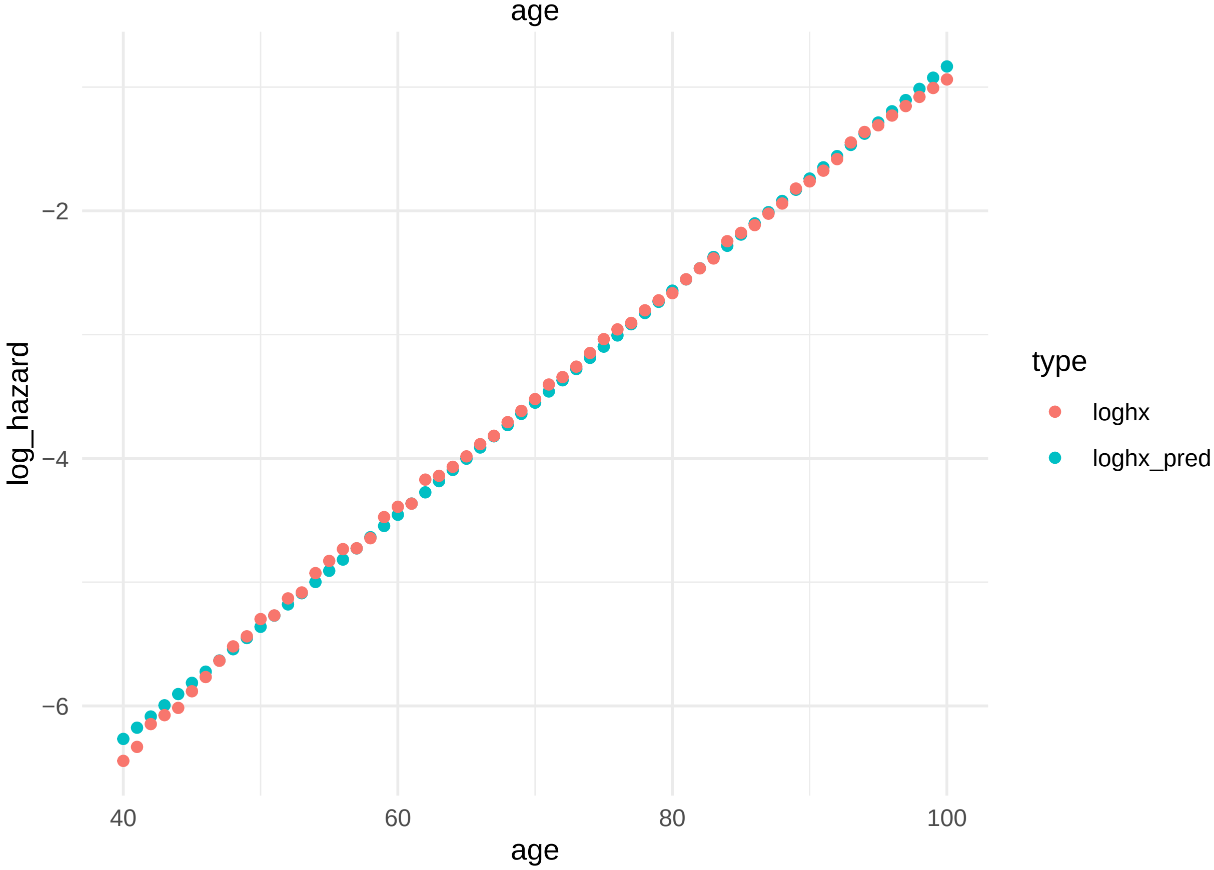


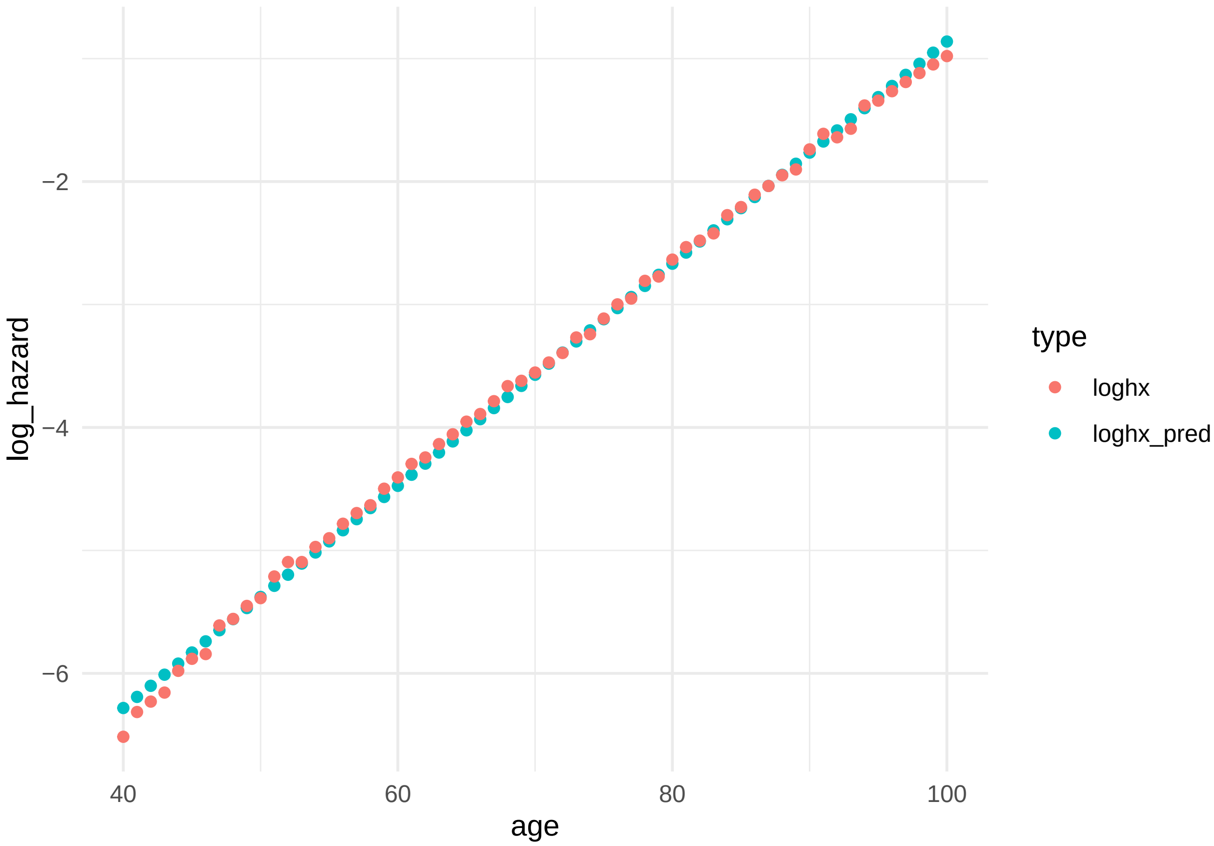
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



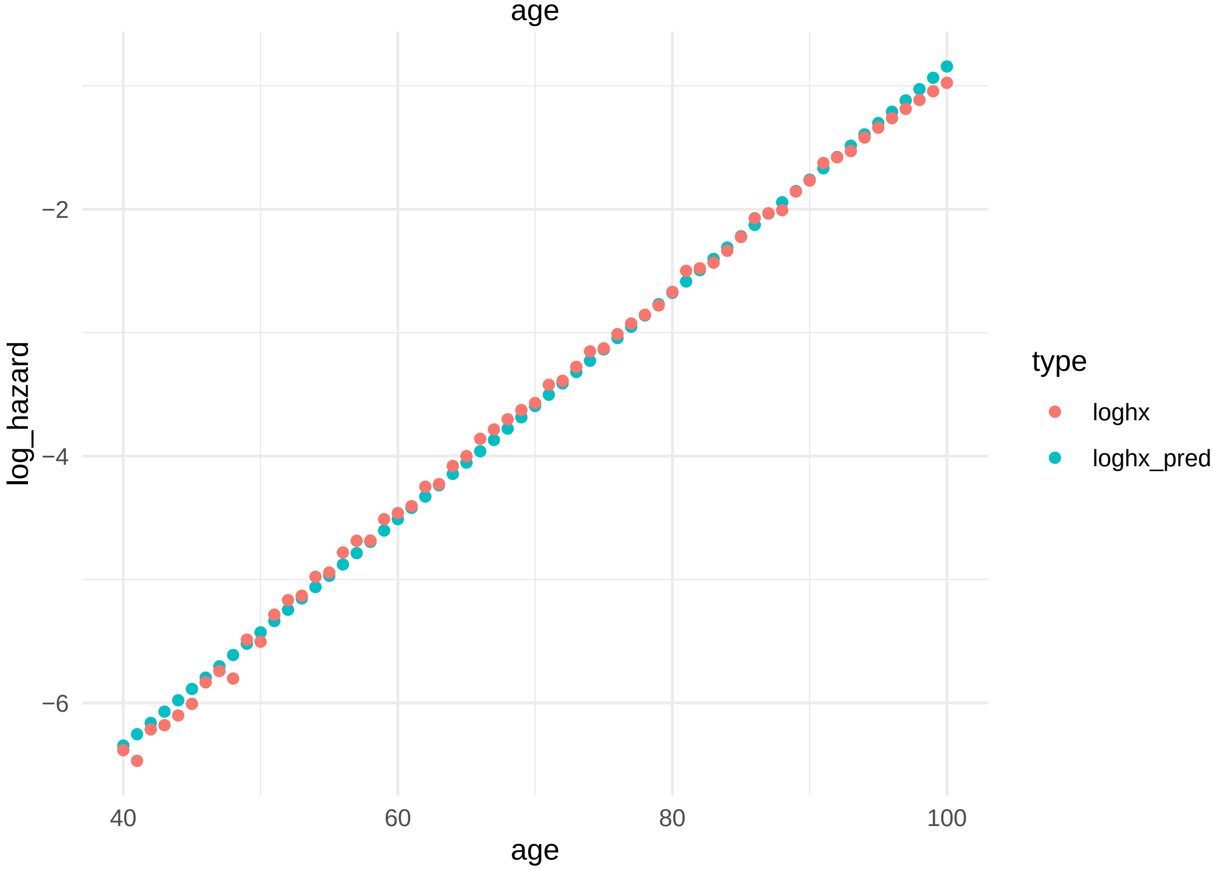
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

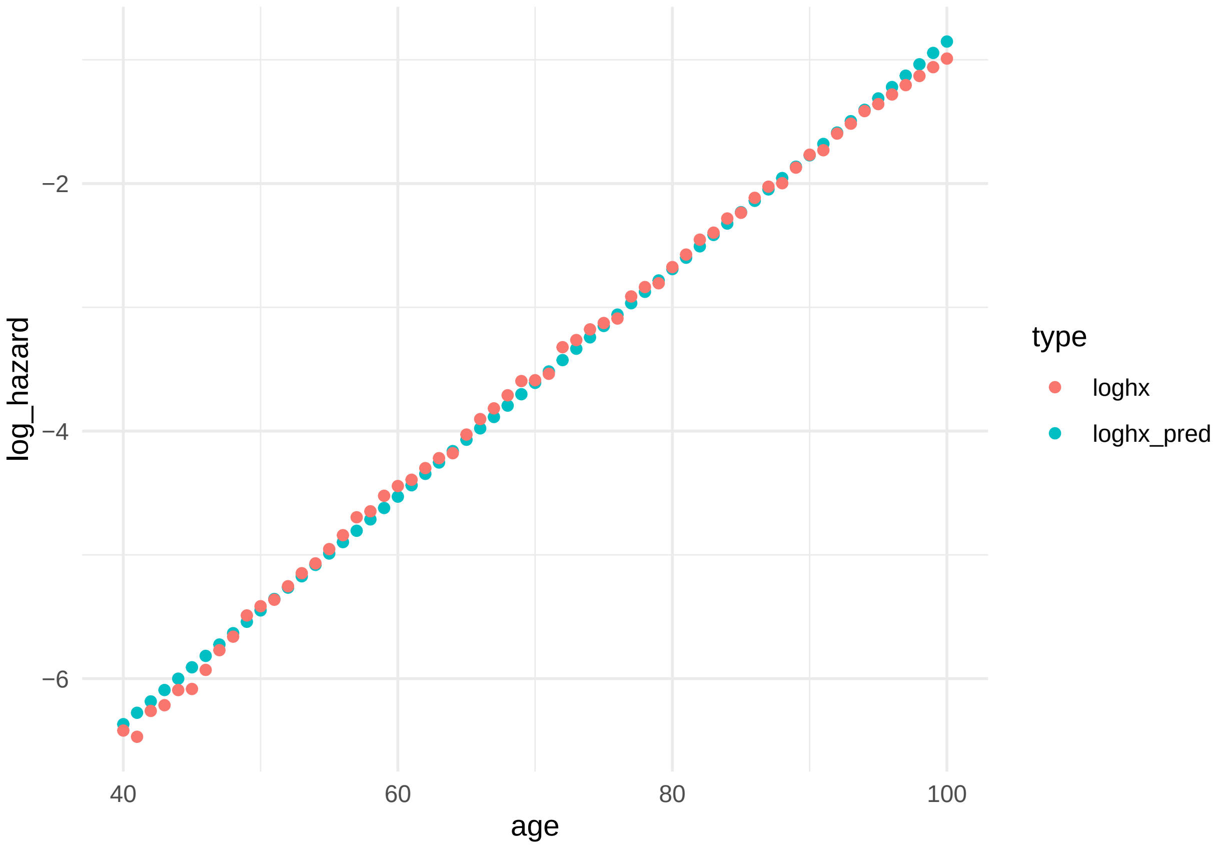

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

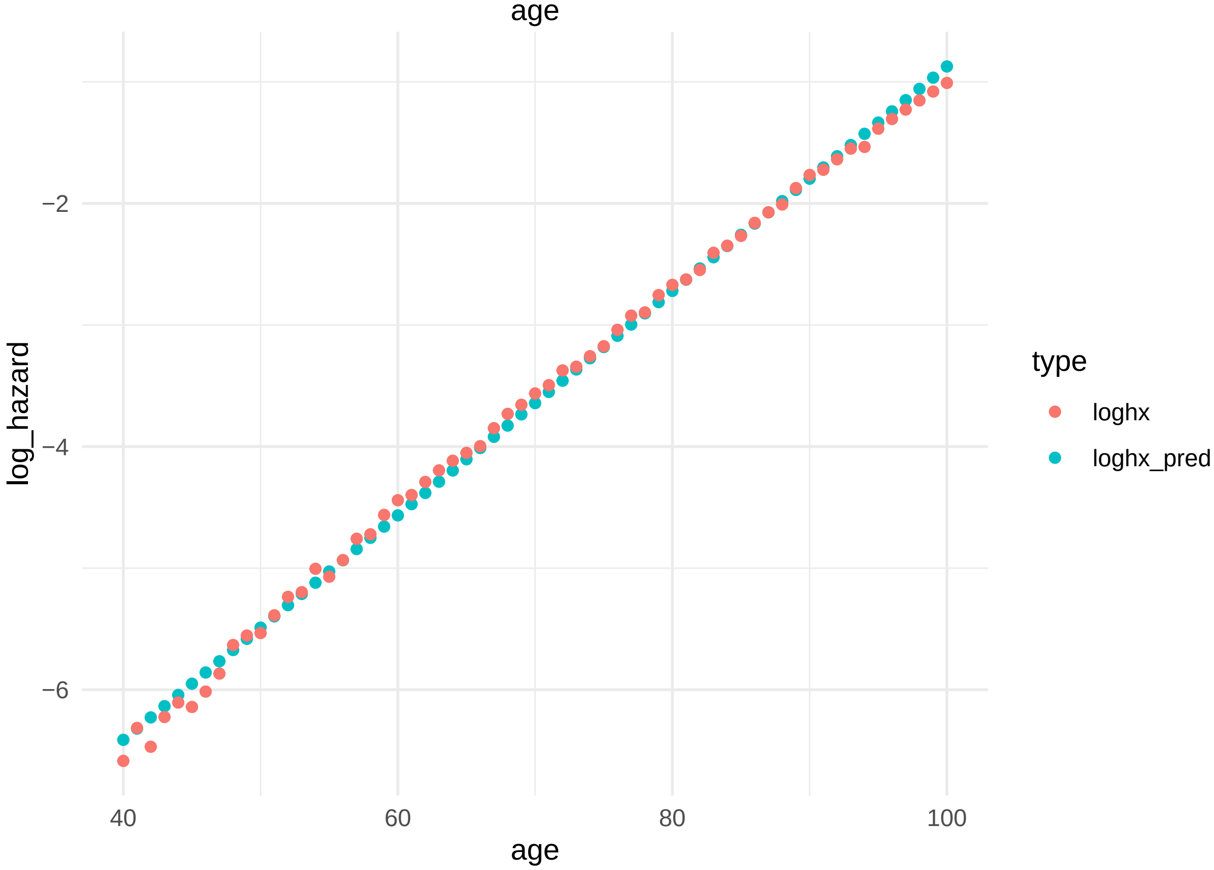


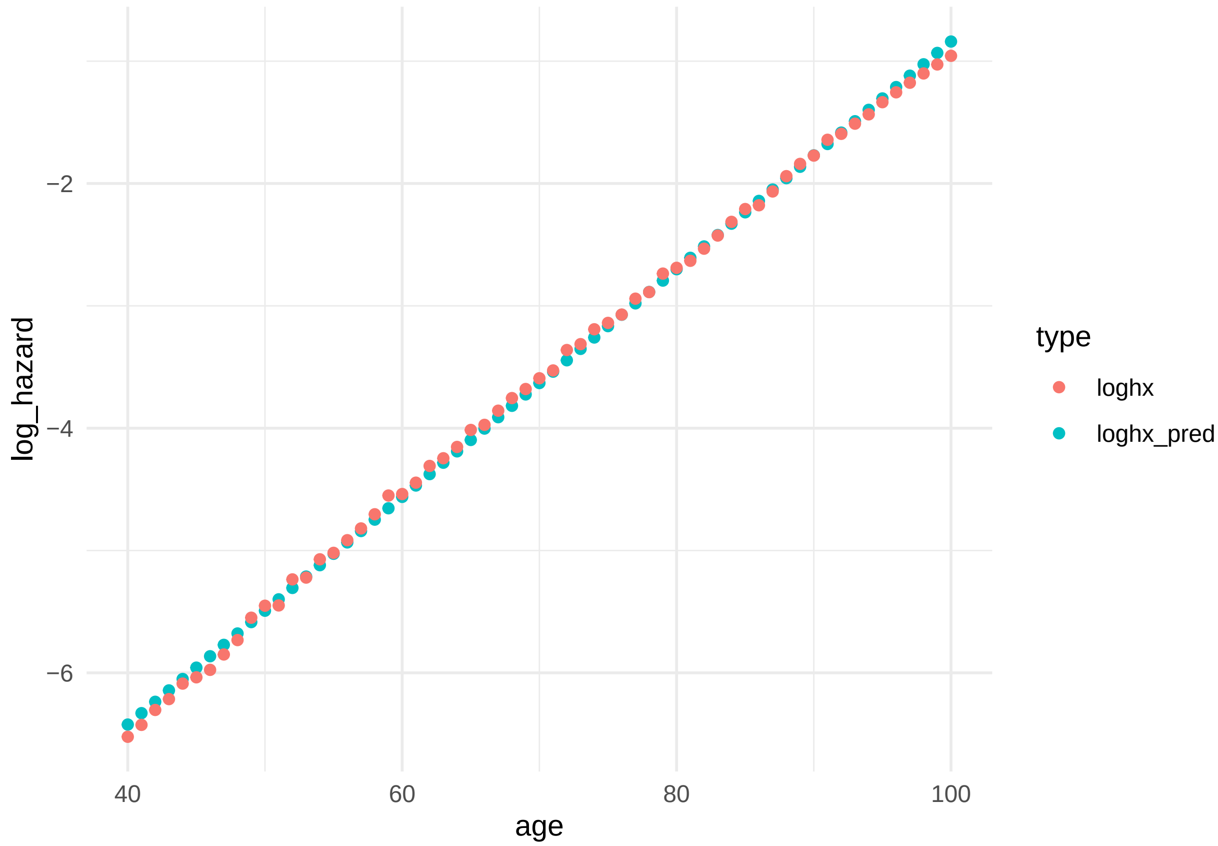
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

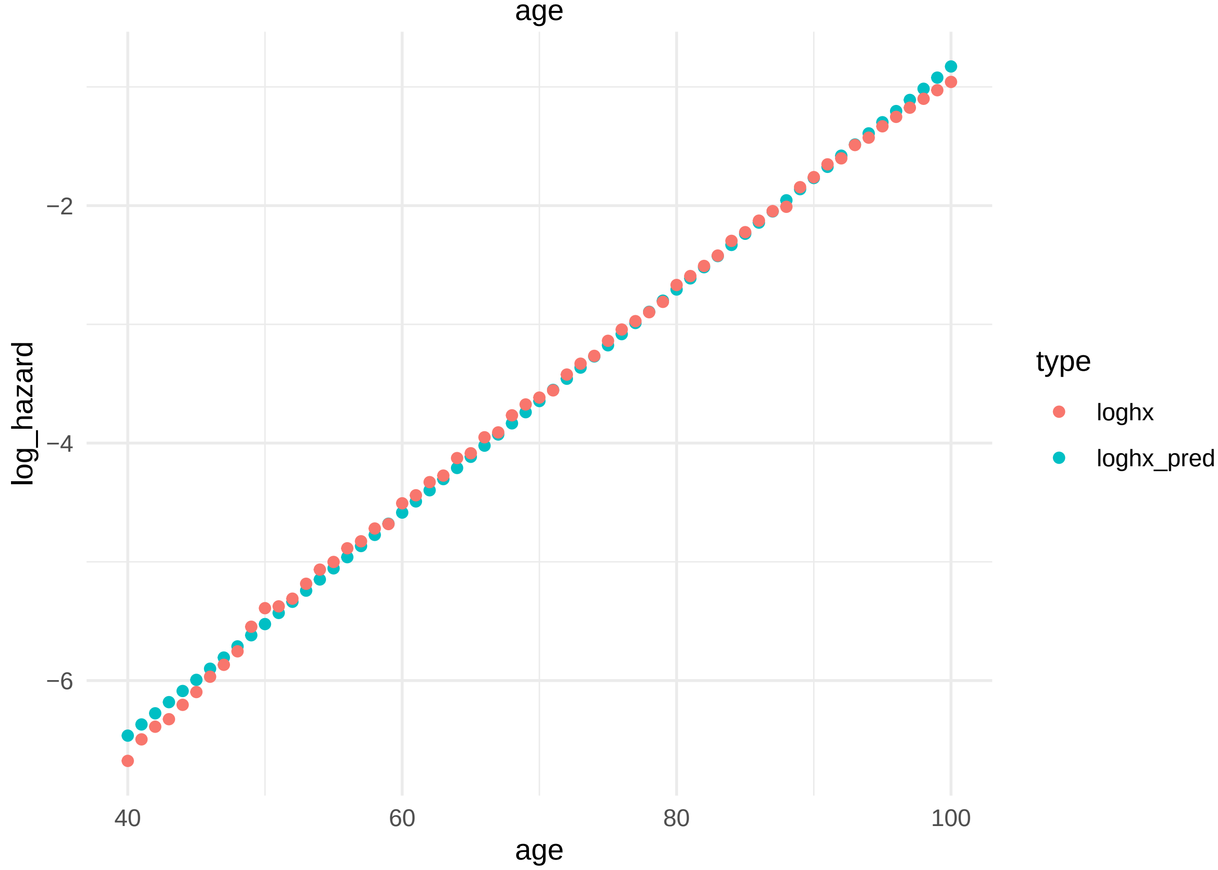


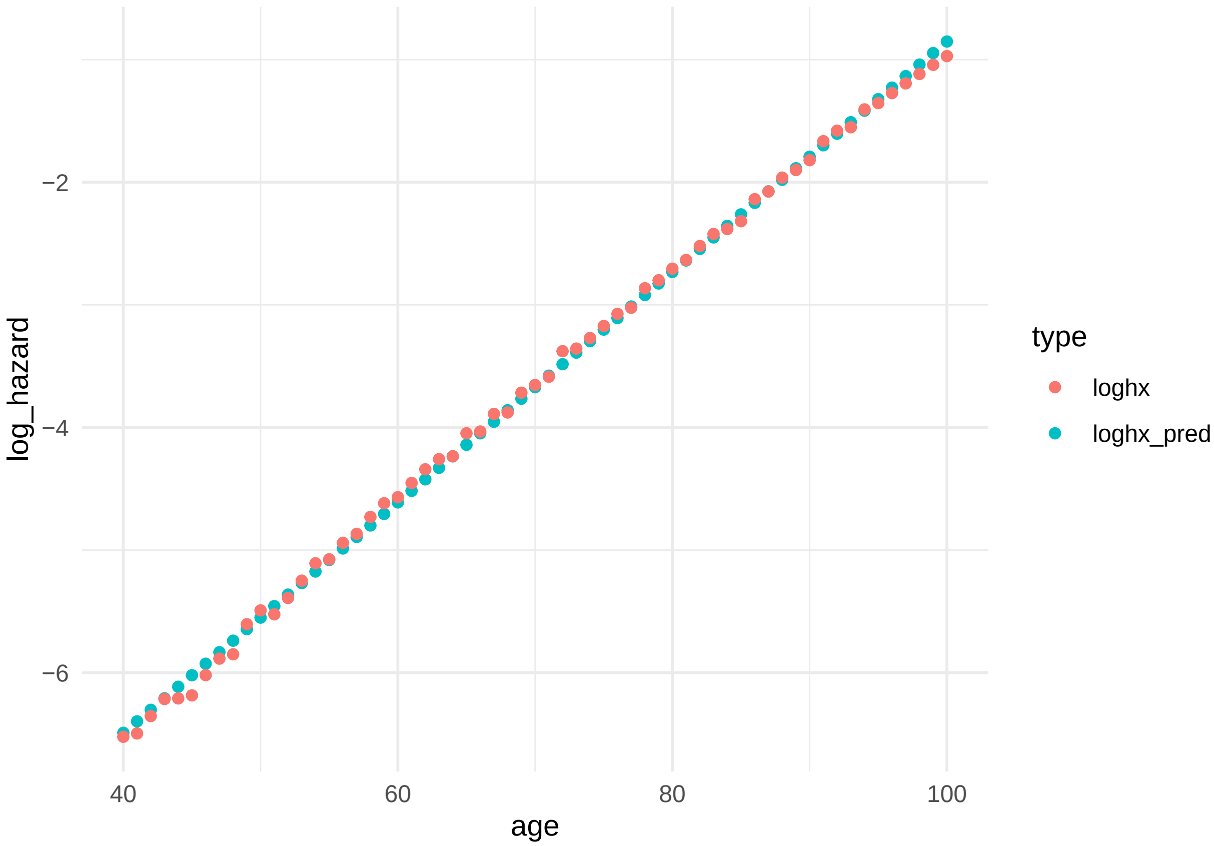
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

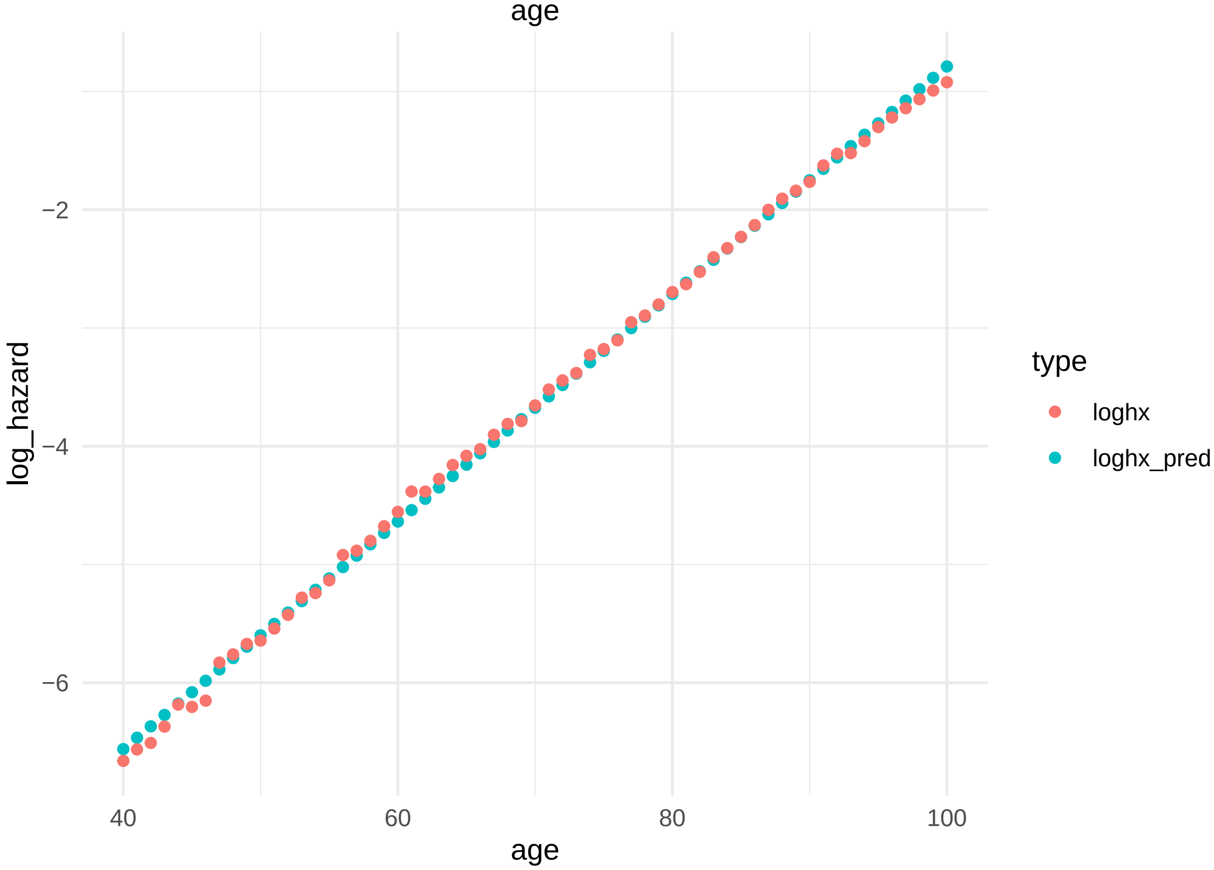

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

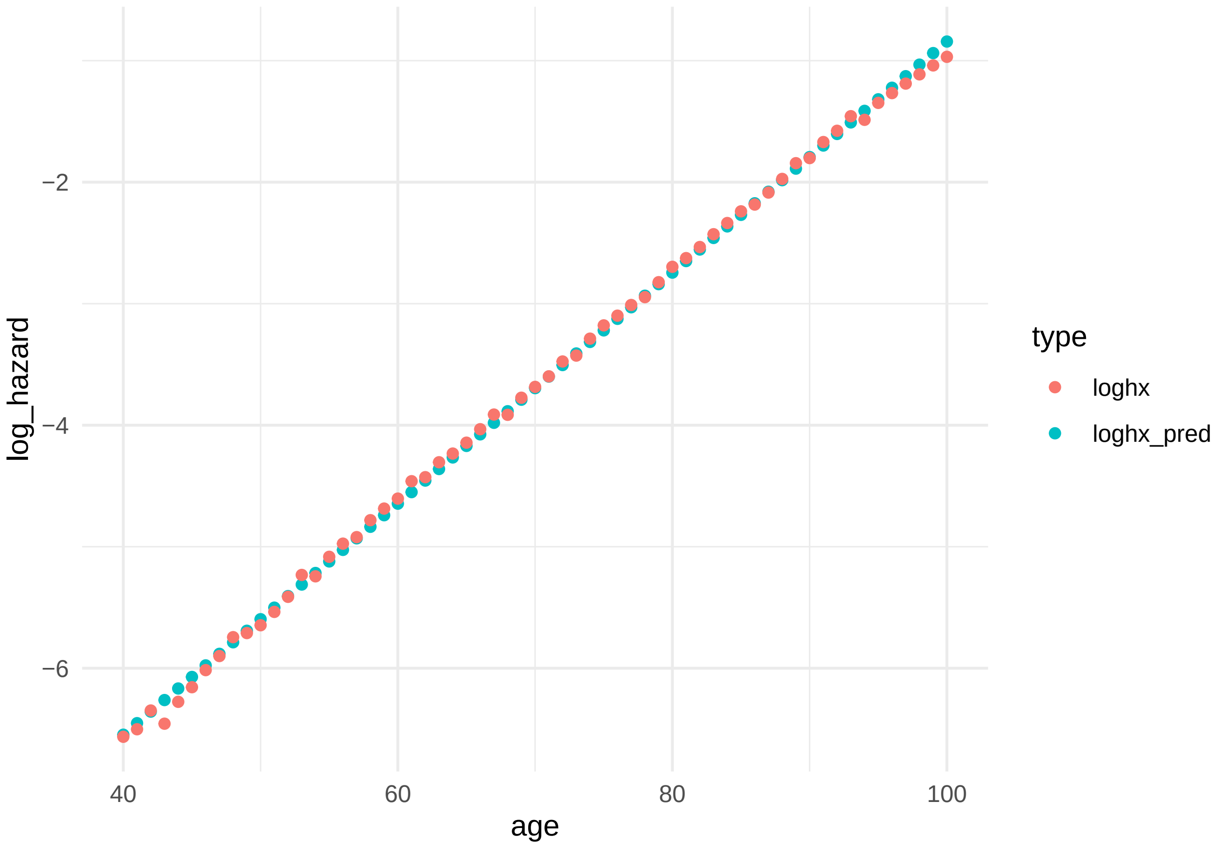

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

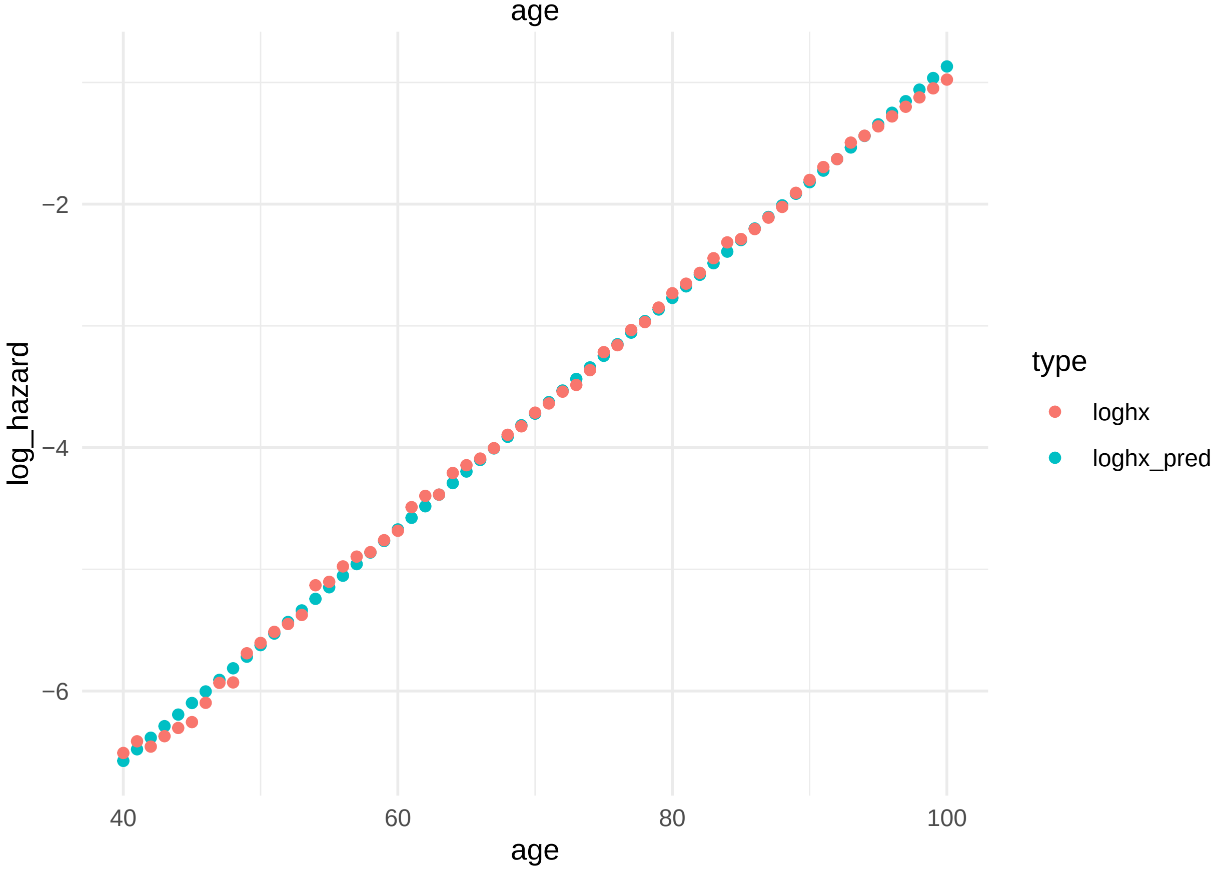

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

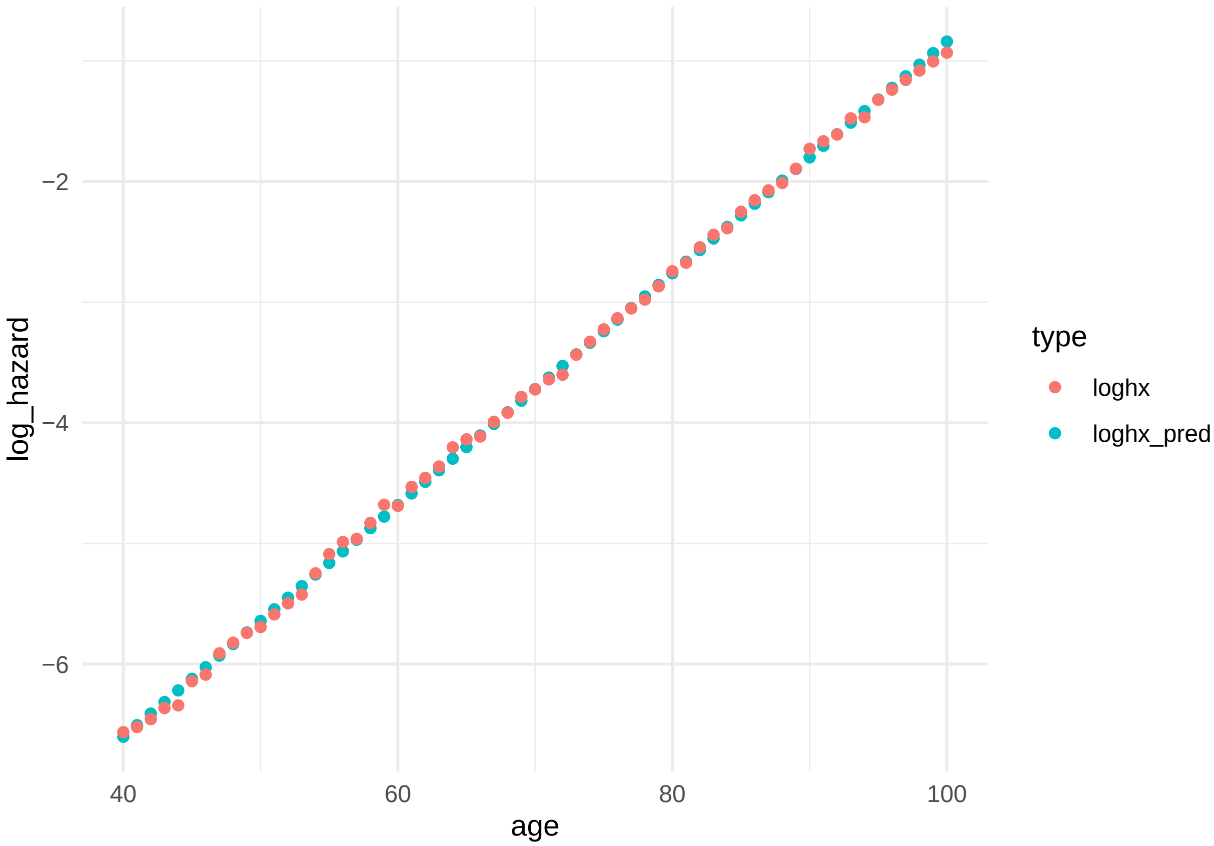

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

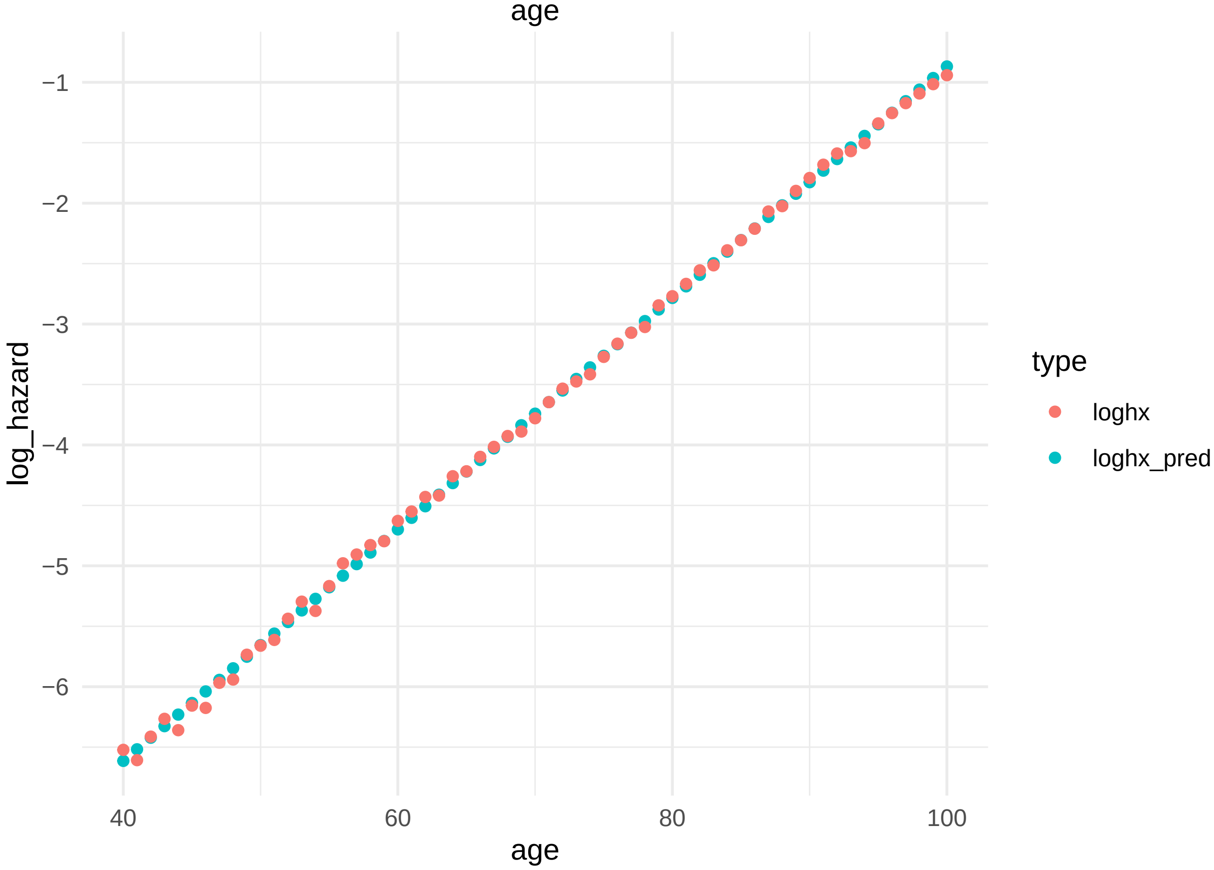

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

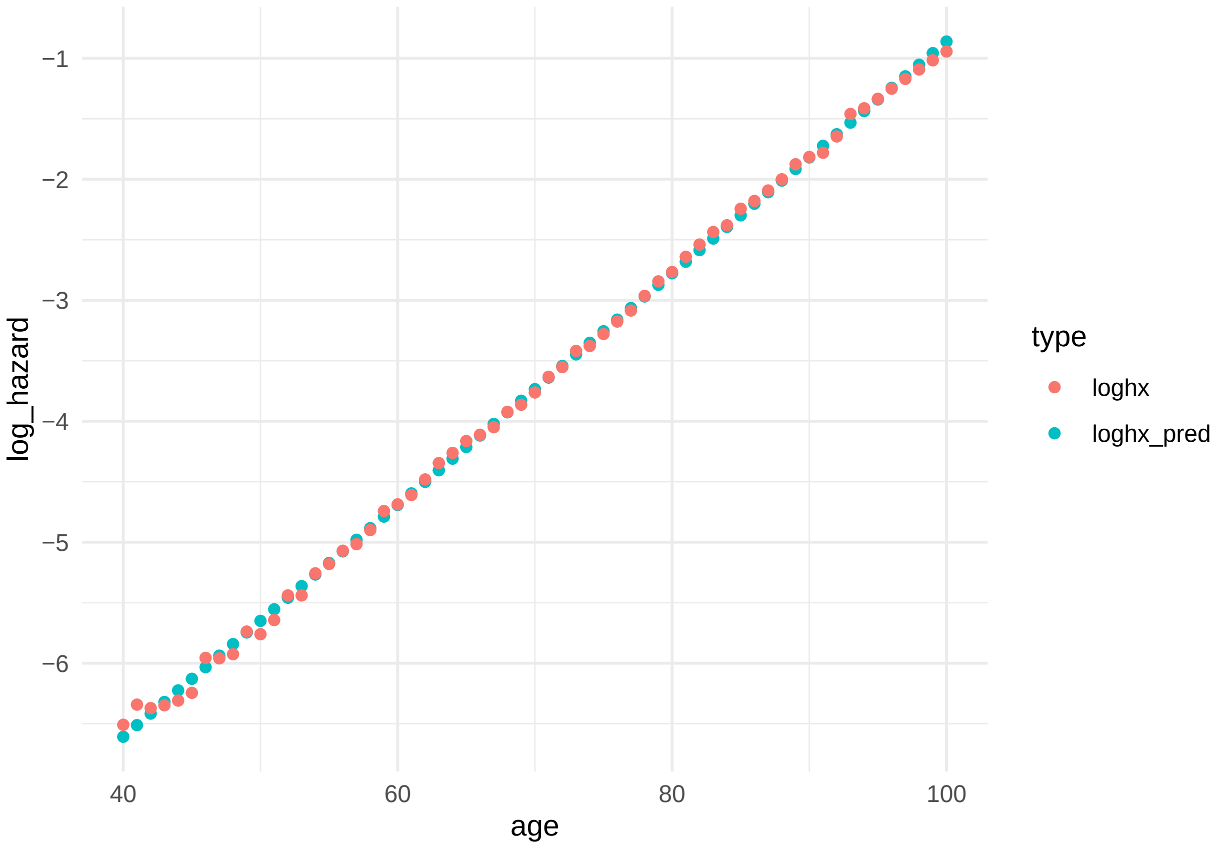

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

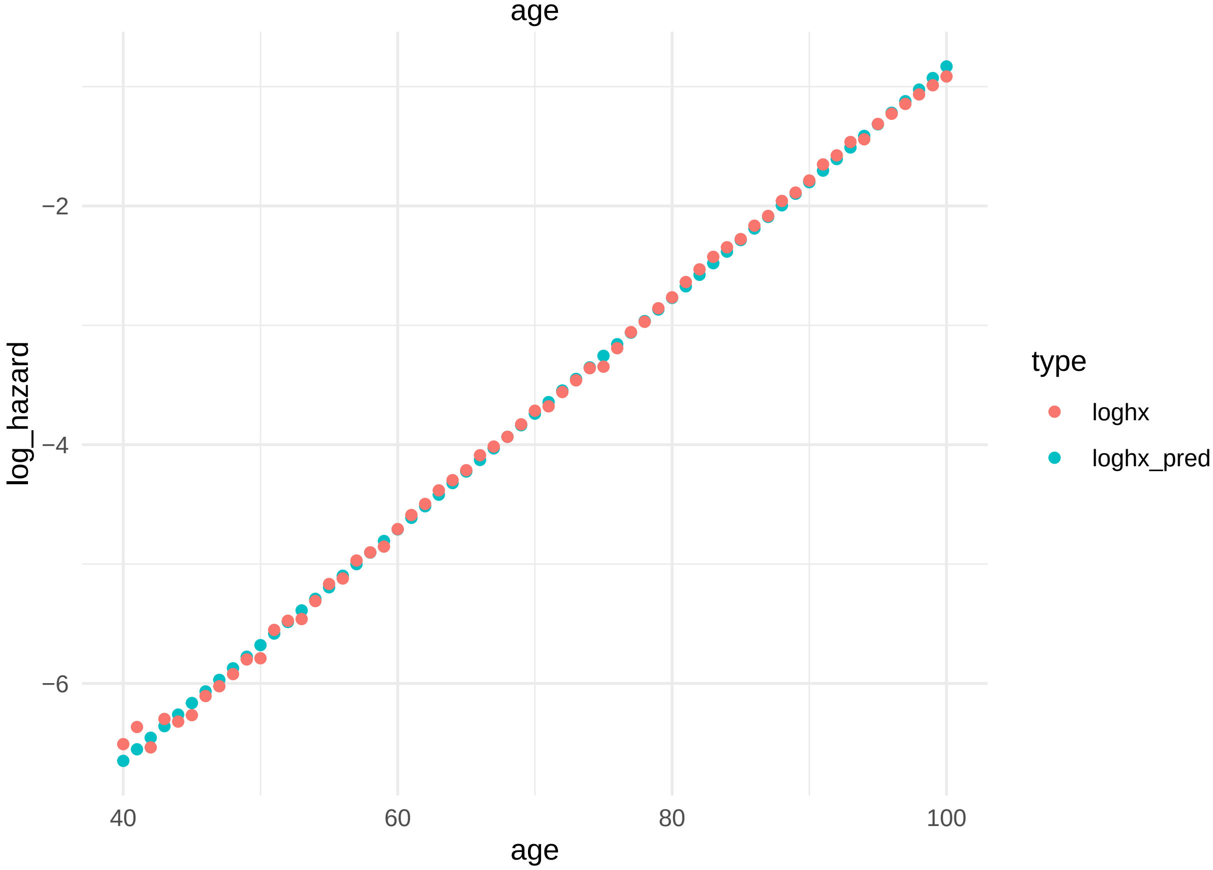

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

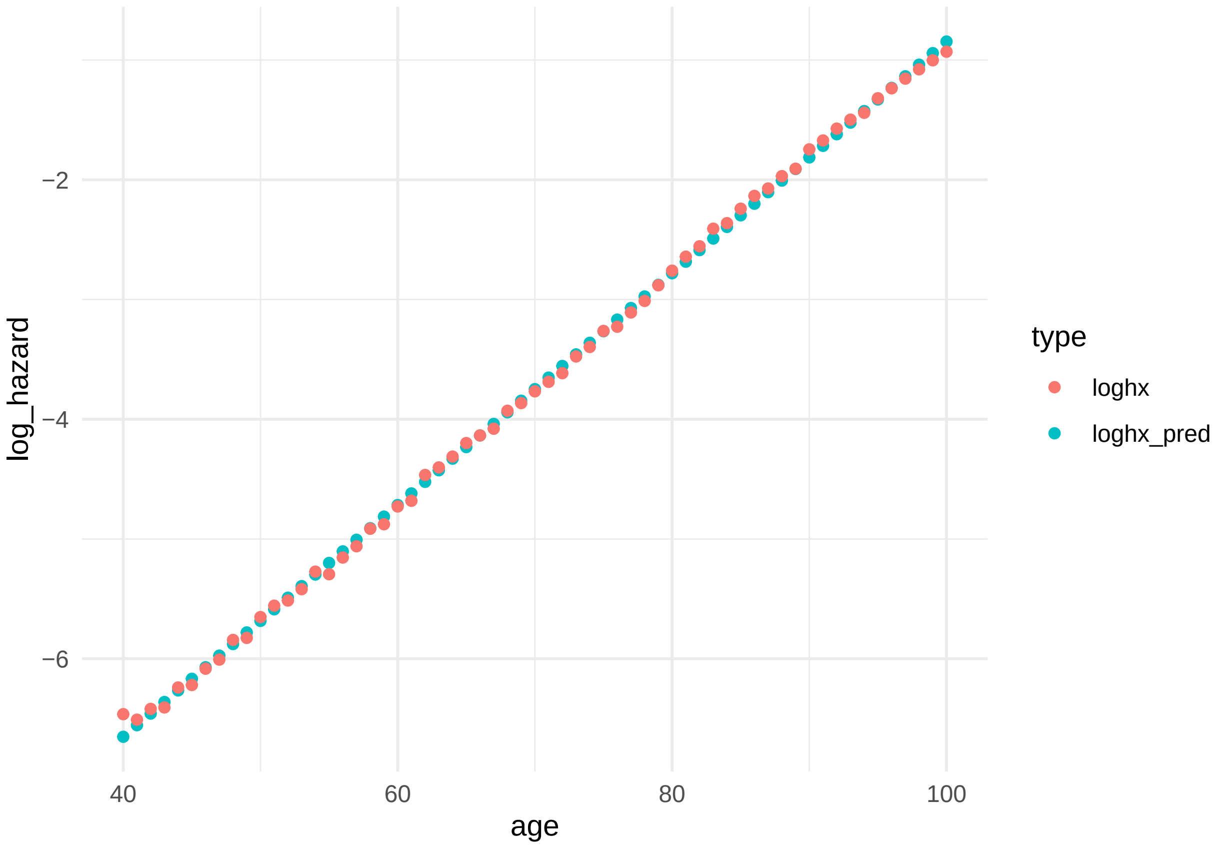

log_hazard
−1
−2
−3
−4
−5
−6
40
60
80
100
age
type
loghx
loghx_pred

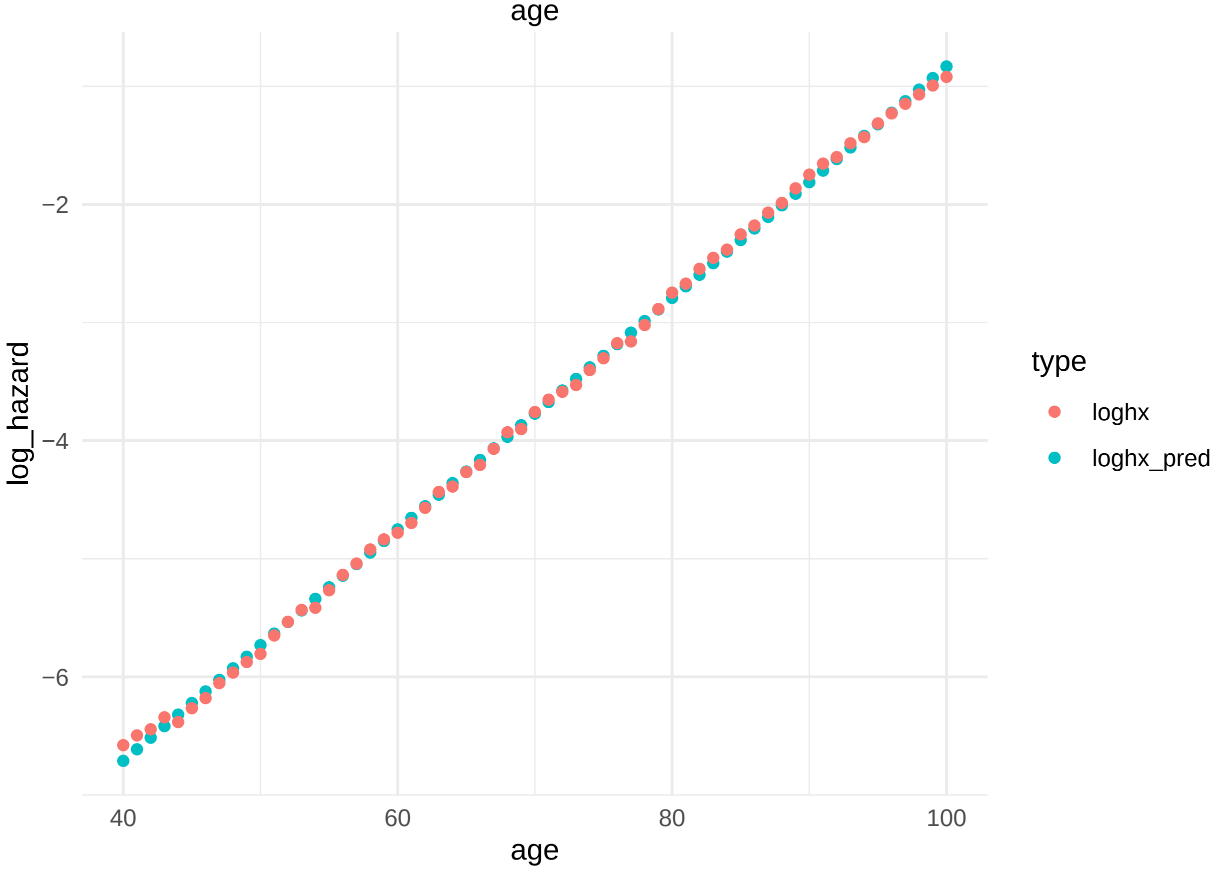


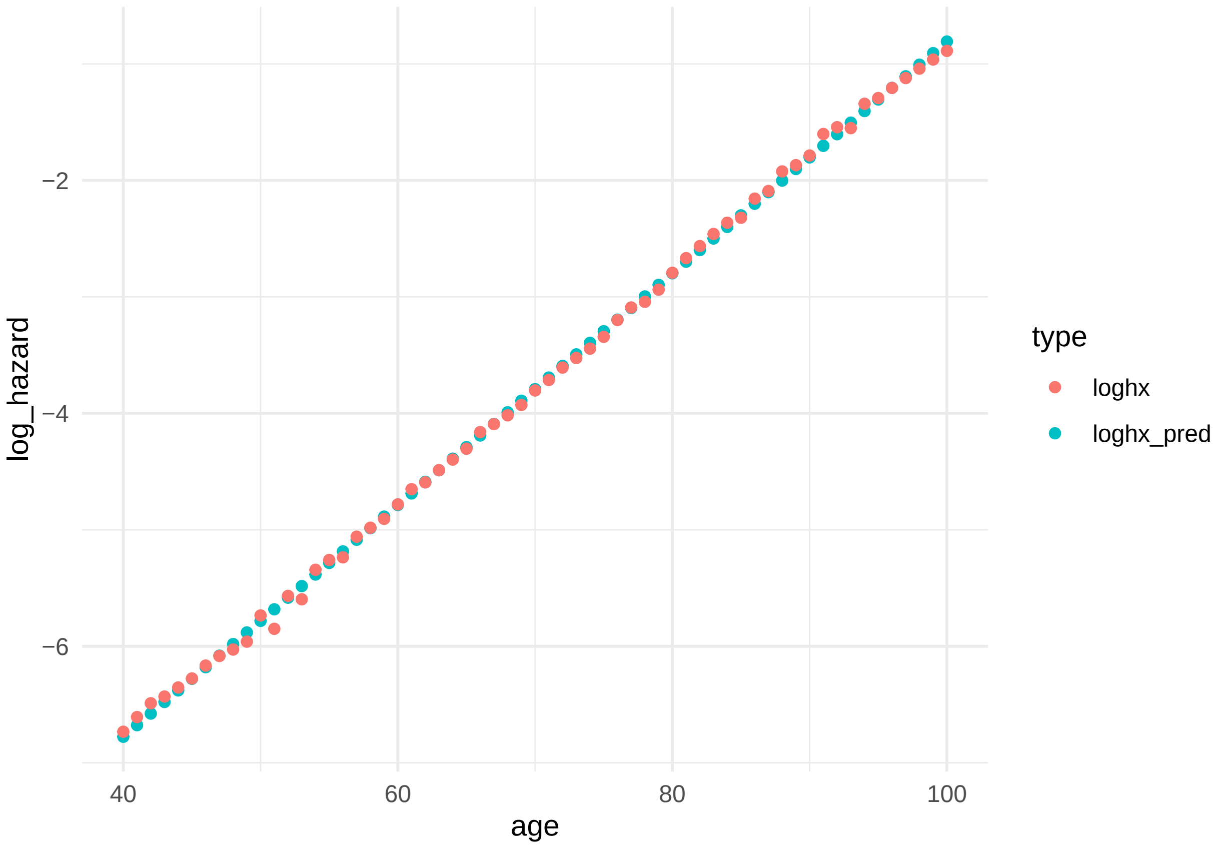
−1
−2
−3
−4
−5
−6
log_hazard
40
60
80
100
age
type
loghx
loghx_pred



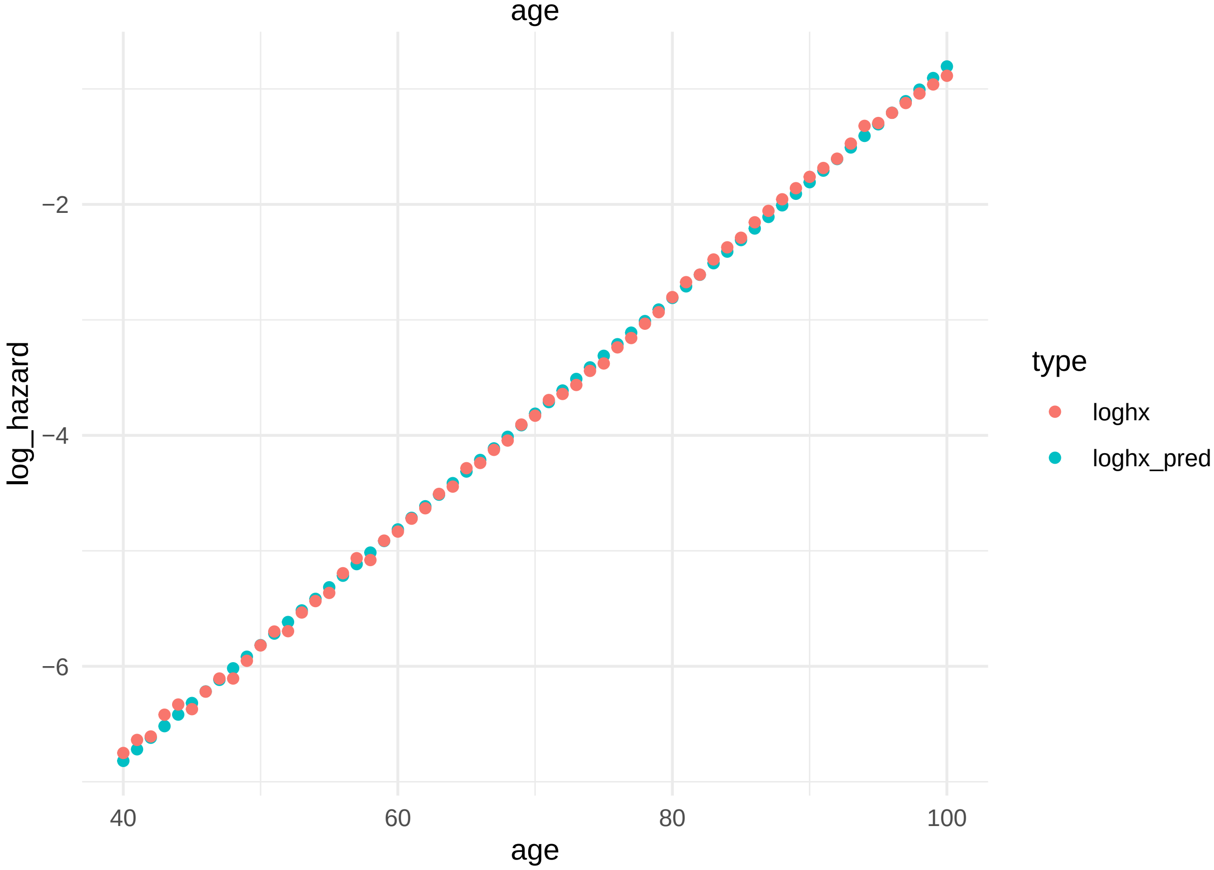
−2
−4
−6
log_hazard
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

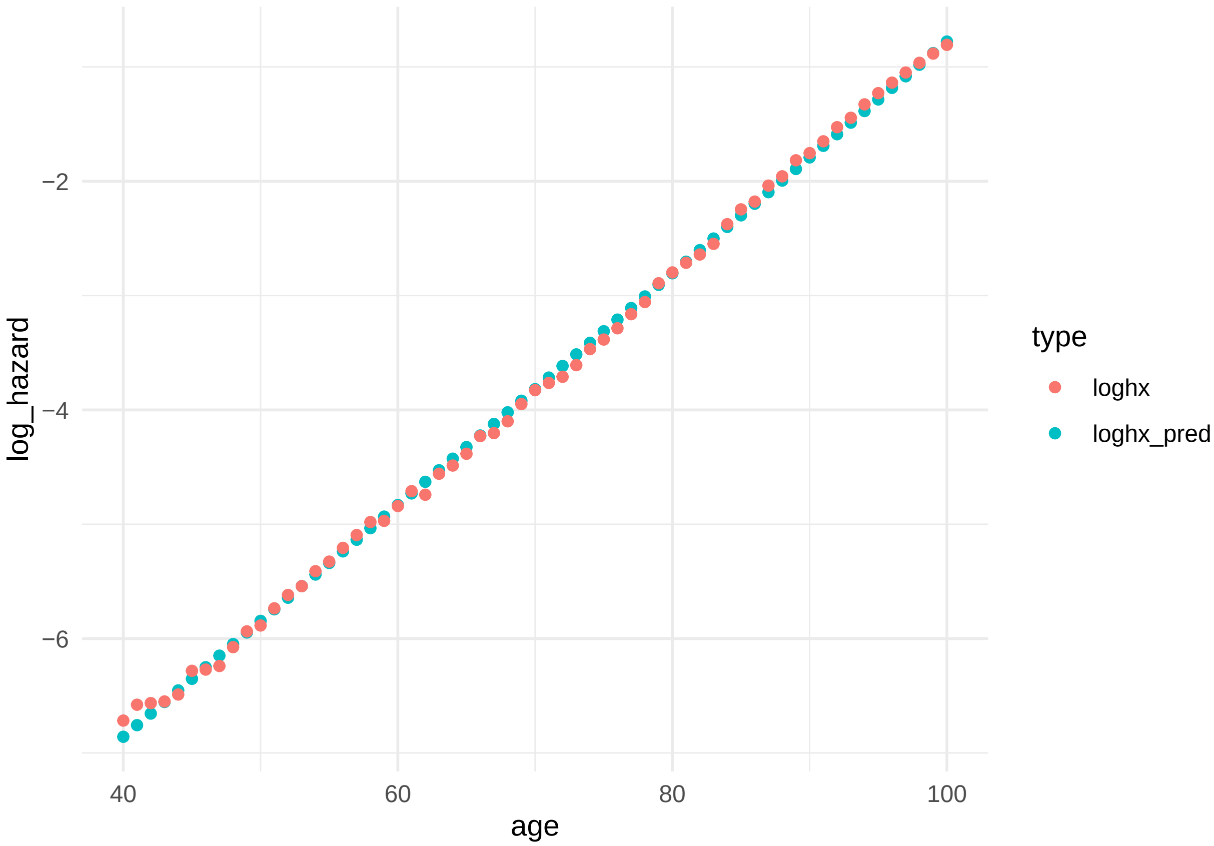

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

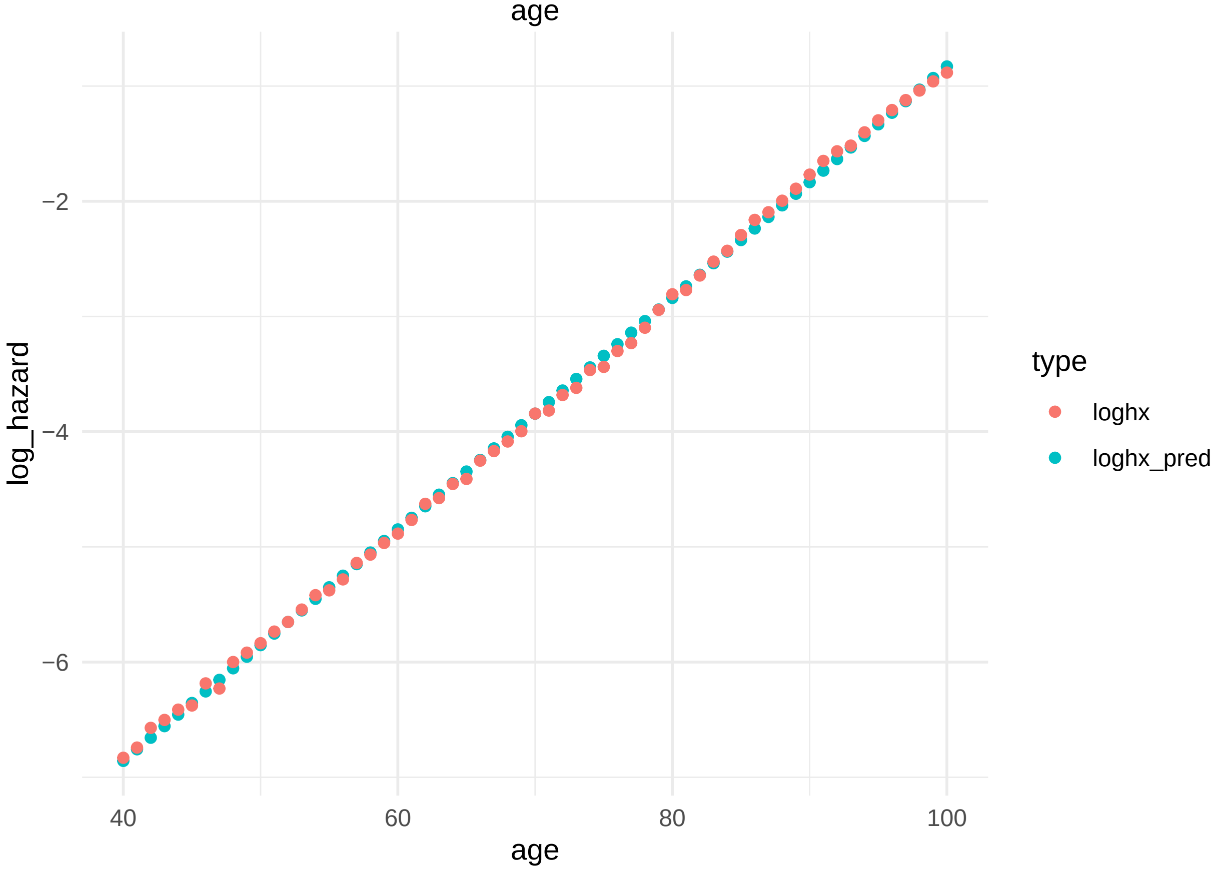


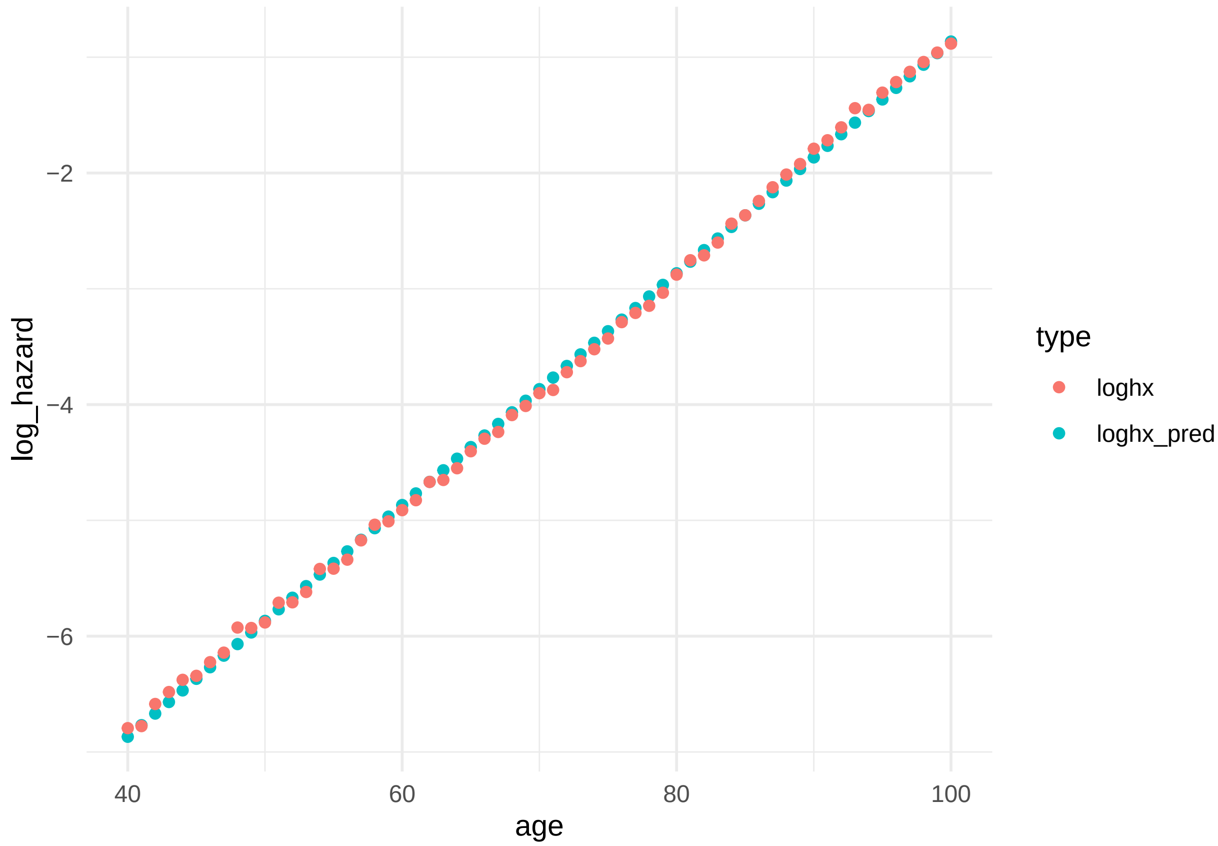
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

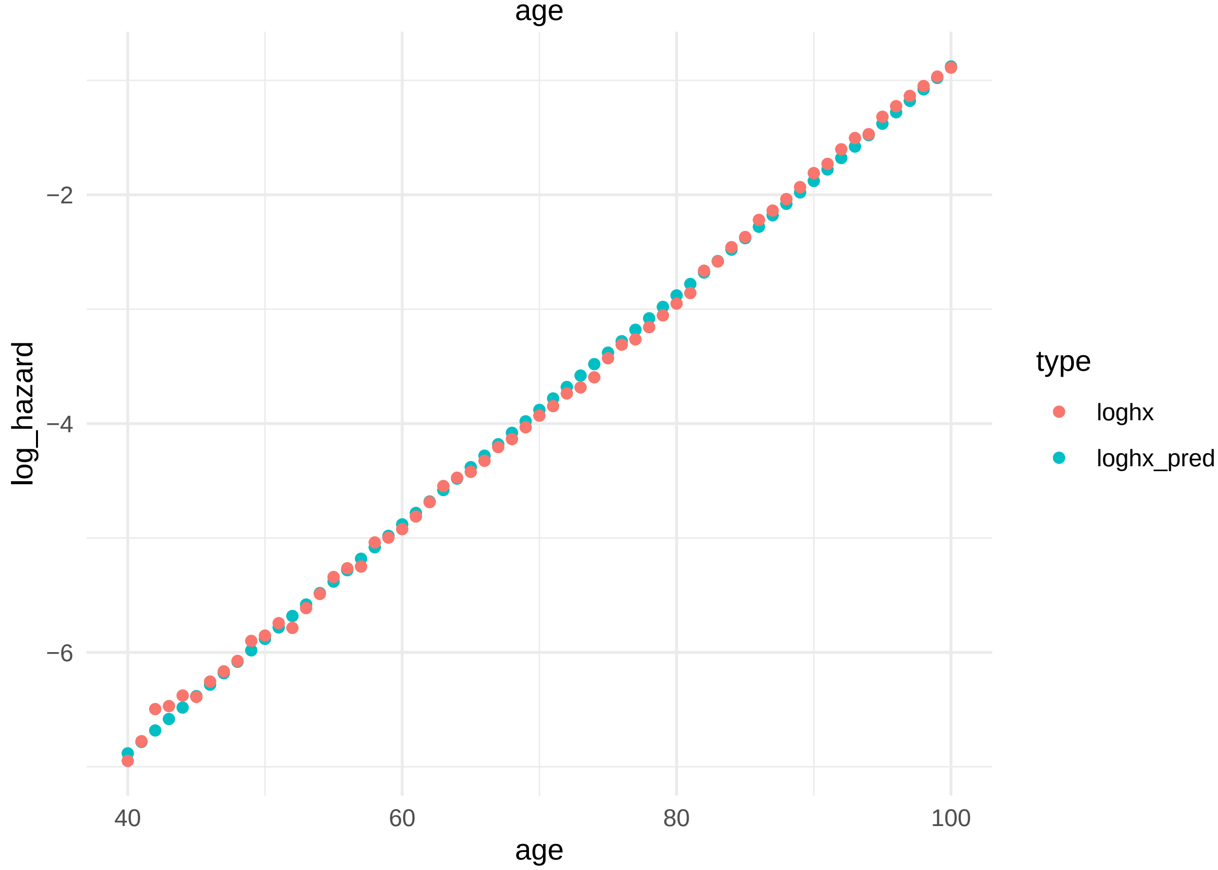


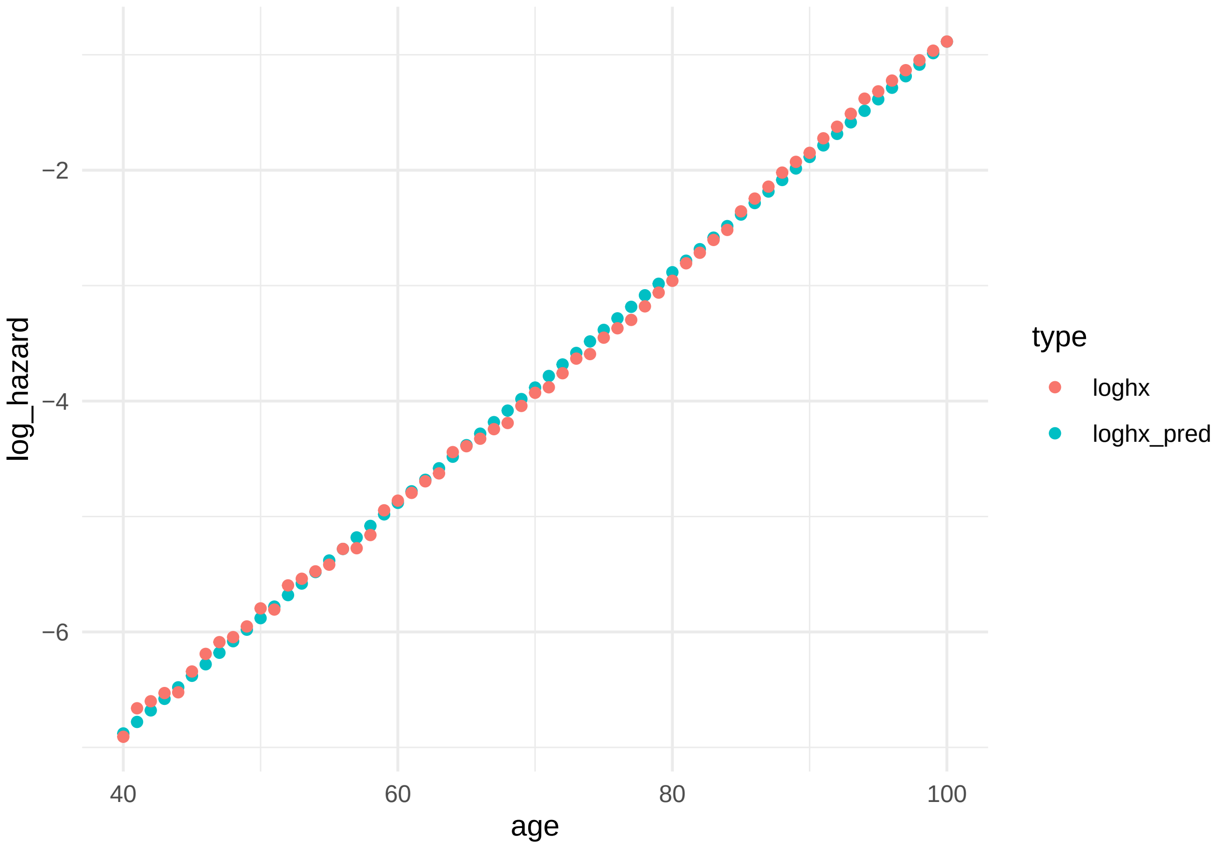
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

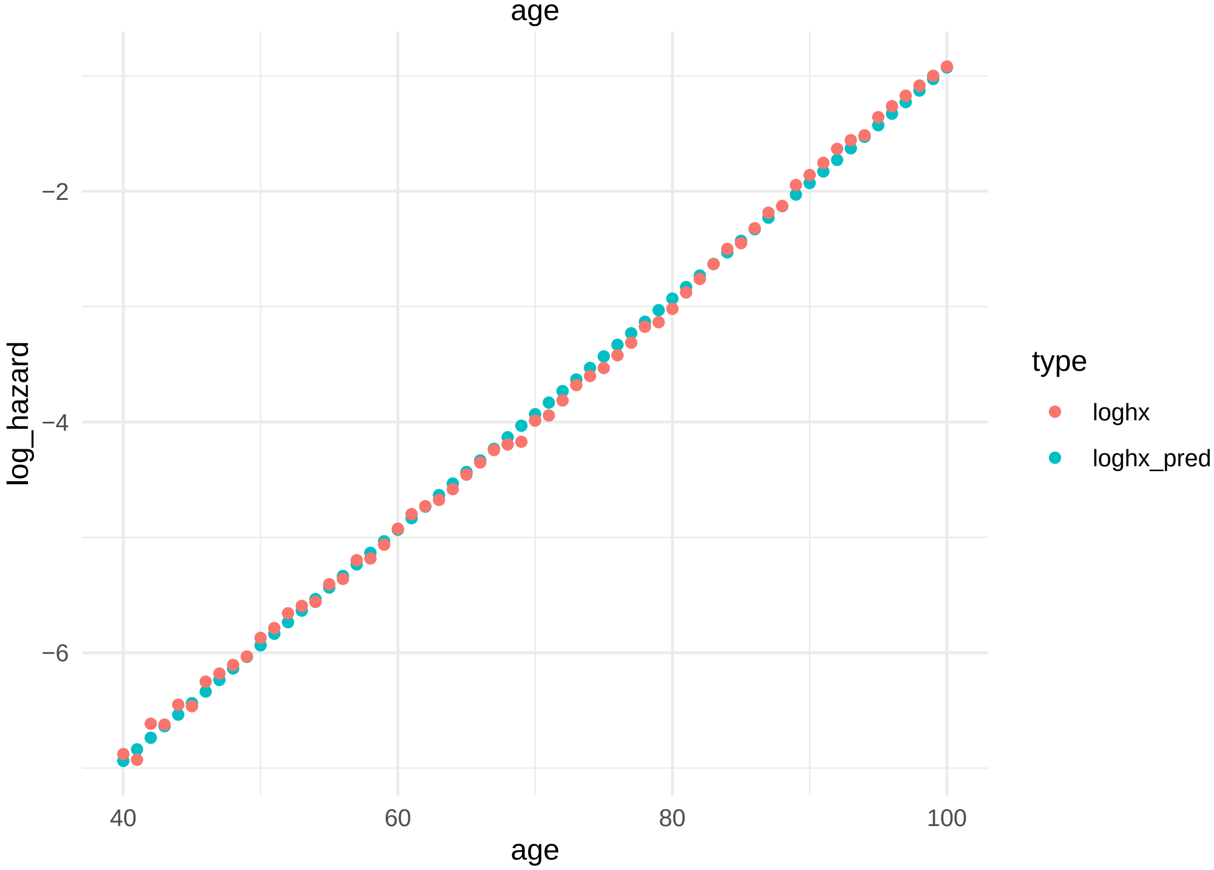

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

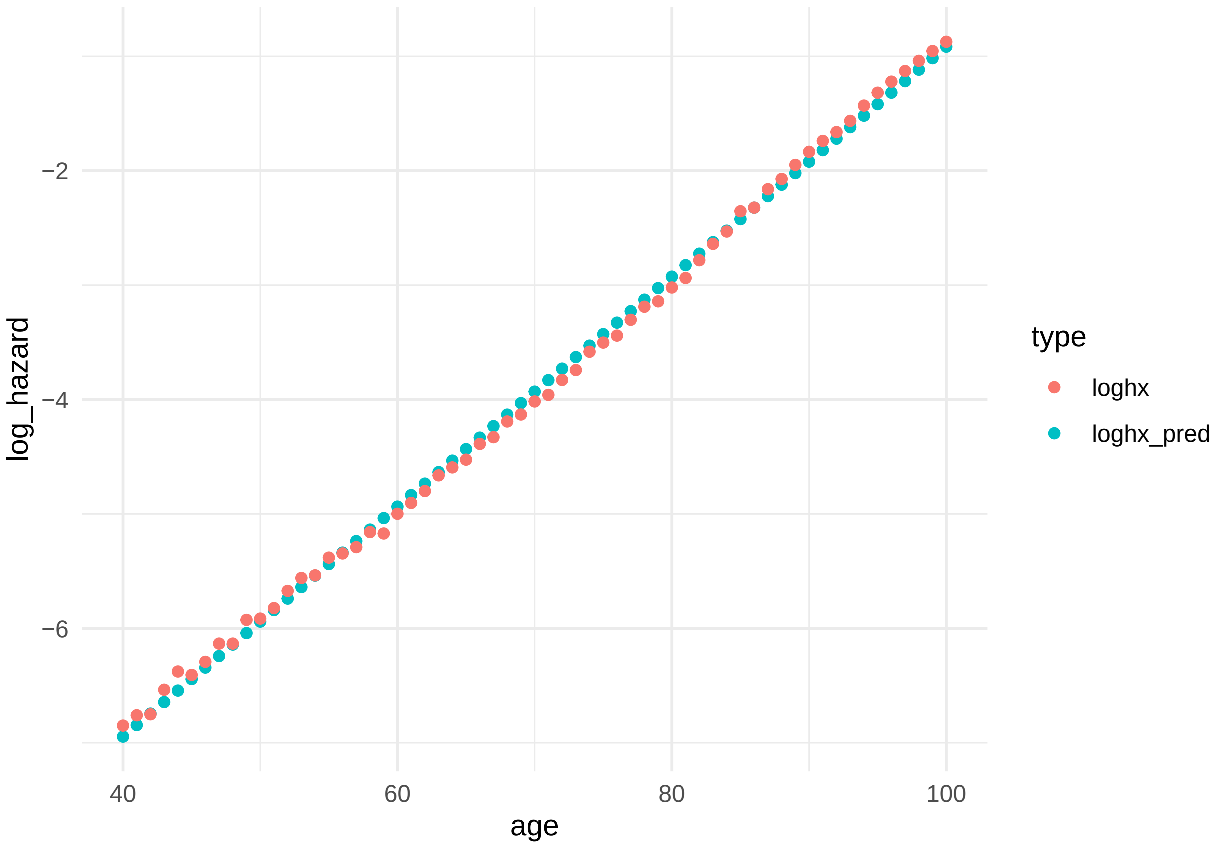

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

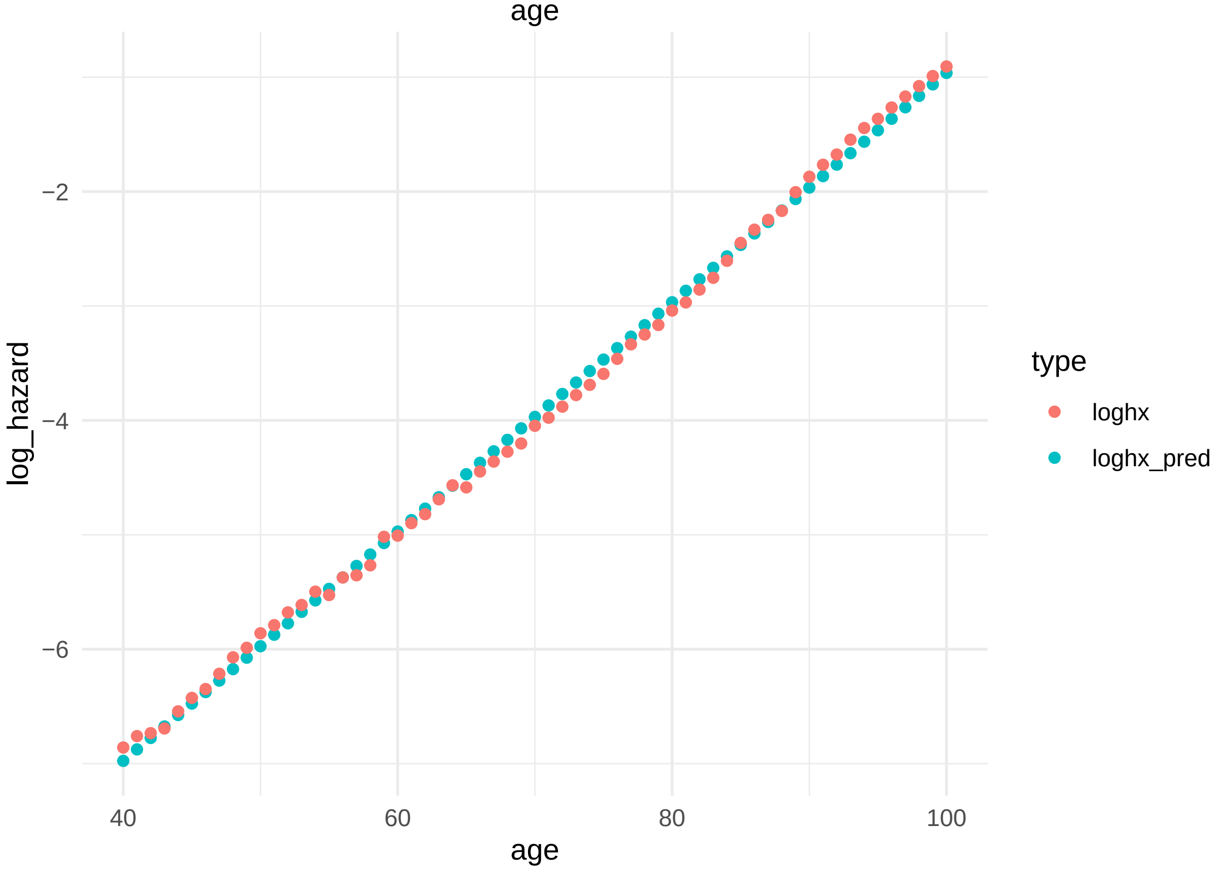


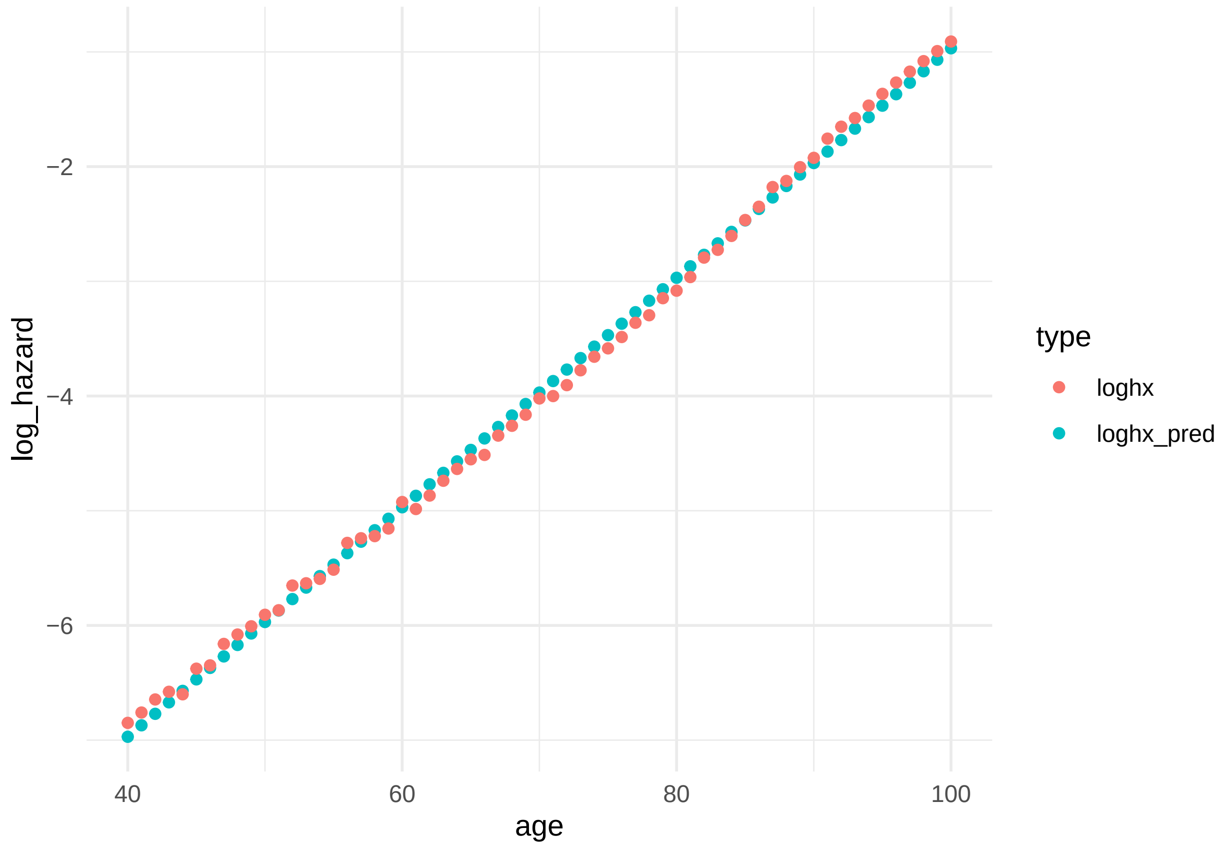
−2
−4
−6
log_hazard
40
60
80
100
age
type
loghx
loghx_pred



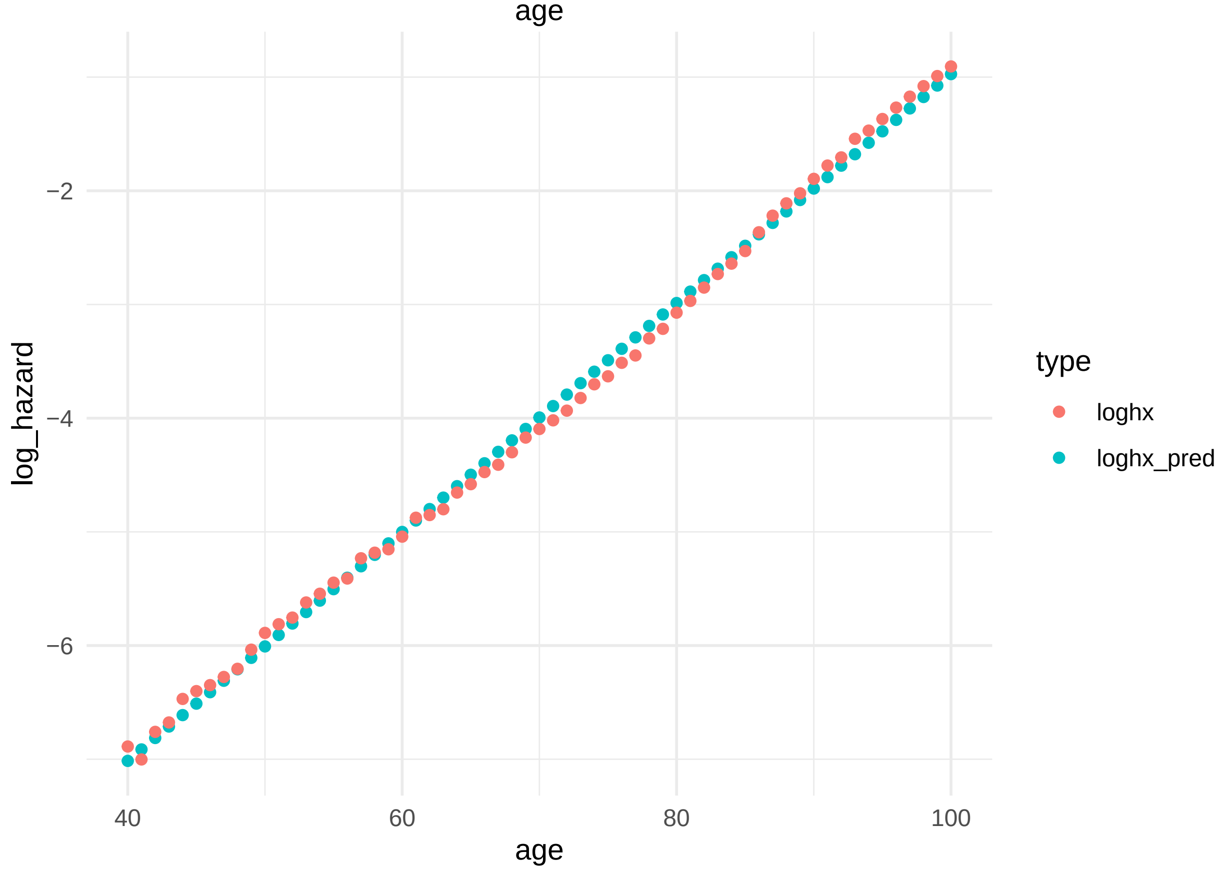
−2
−4
−6
log_hazard
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

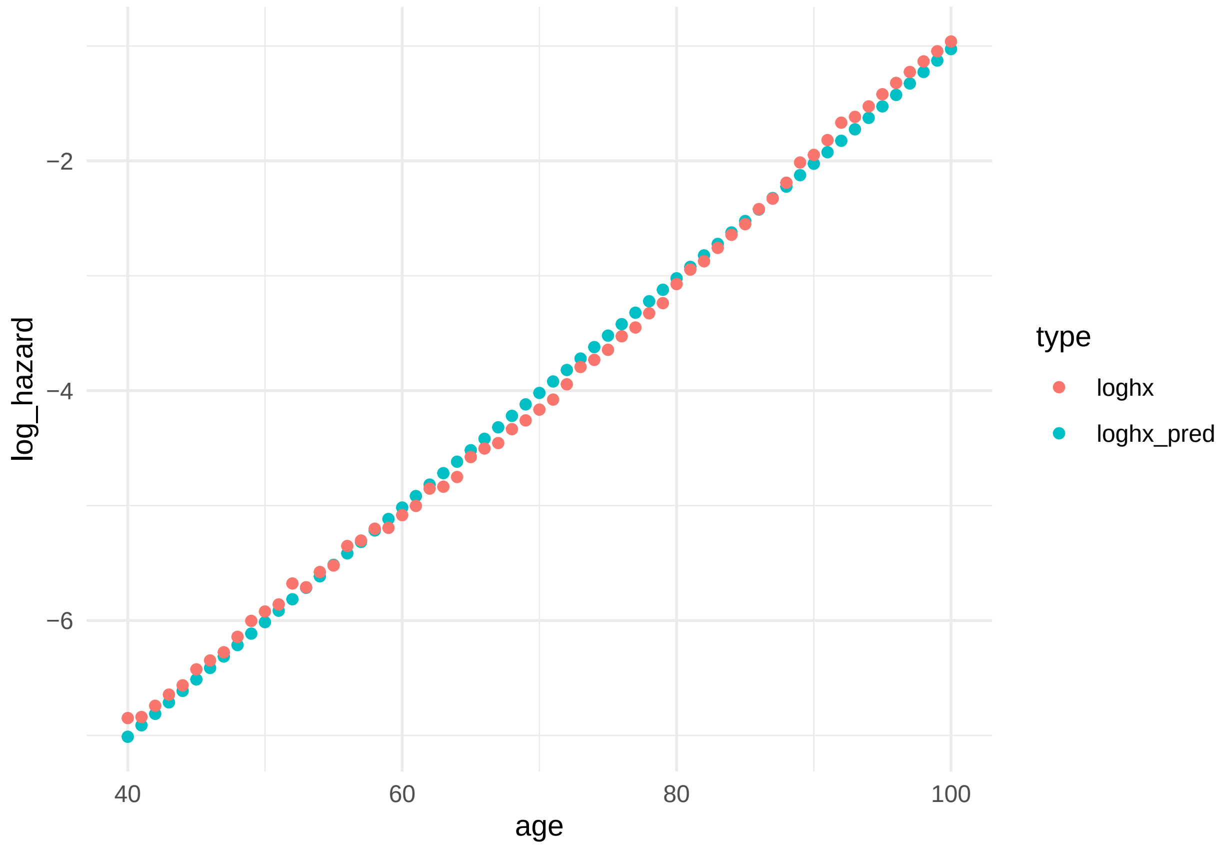

log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred

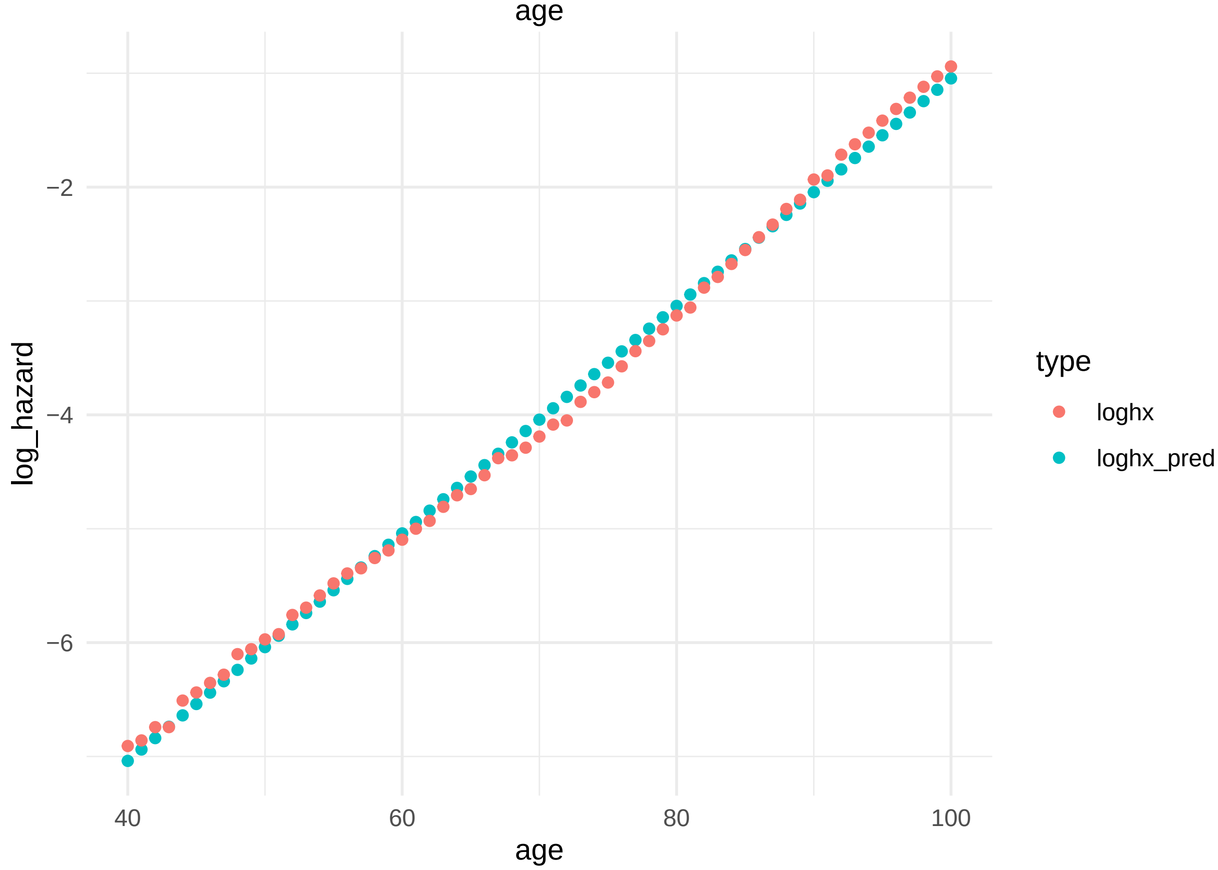


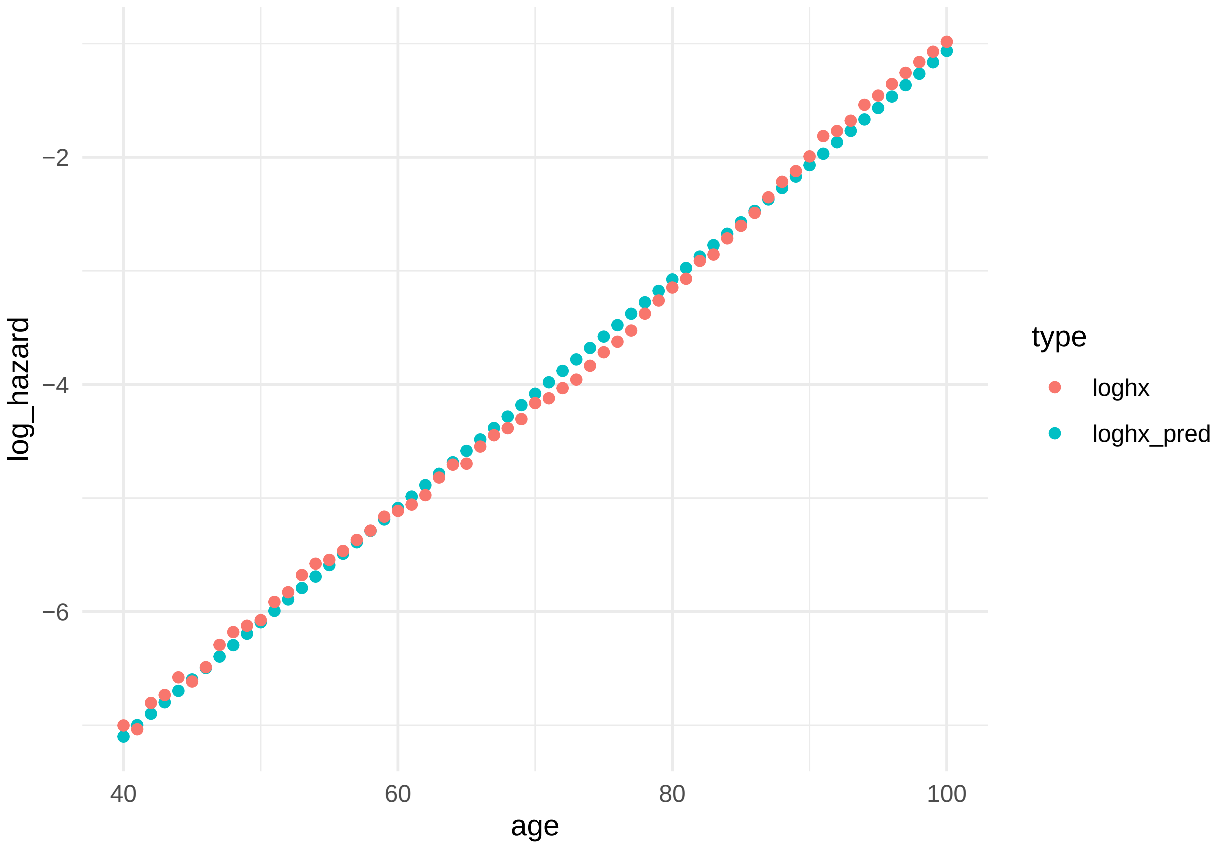
log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred


log_hazard
−2
−4
−6
40
60
80
100
age
type
loghx
loghx_pred



Question 2

a

b

c

d

e

f

g