Censoring_example

Anton Holm

2020-04-28

```
set.seed(19931031)
library(gridExtra)
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::combine() masks gridExtra::combine()
## x dplyr::filter()  masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(latex2exp)

Slope <- log(1.3)
Error_term <- rlnorm(1000, 0, 0.1)
Covariate <- rnorm(1000, 0, 1)
Response <- exp(Covariate*Slope)*Error_term
LOD <- sort(Error_term)[400]
Censored_data <- ifelse(Response<LOD, -LOD, Response)
data <- as.data.frame(Censored_data) %>%
  mutate(cens = Censored_data < 0) %>%
  mutate(Covariates = Covariate) %>%
  mutate(Censored_data = ifelse(Censored_data < 0, log(abs(Censored_data)/sqrt(2)), log(abs(Censored_da
  mutate(cens = ifelse(cens == FALSE, "Uncensored", "Censored")) %>%
  mutate(true_value = log(Response))

censored_plot <- data %>% ggplot(aes(x=Covariates, y = Censored_data, shape = cens)) +
  geom_point(alpha = 0.3) +
  scale_y_continuous(limits = c(-0.9, 1)) +
  geom_hline(yintercept = log(LOD), linetype = 'dashed') +
  scale_shape_manual(values = c(16,2)) +
  xlim(-4, 4) +
  theme_minimal() +
  theme(plot.caption=element_text(hjust = 0), legend.position = c(0.77, 0.15), legend.background = eleme
    fill = NA, size =0.5, linetype = "solid"),
        panel.border = element_rect(colour = "black", fill=NA, size=0.5),legend.title = element_blank()
  labs(x='X', y = 'Log(Y)', caption = 'Figure 1a: Example of 1000 observations following a
log-normal distribution. Censored observations are
replaced with the LOQ divided by the squareroot of 2') +
  geom_text(aes(x = 3.5, y = 0.05), label = "LOQ", size = 3, show.legend = FALSE, family = "serif") +
  labs(x = 'X', y = 'Y')
```

```
uncensored_plot <- data %>% ggplot(aes(x=Covariates, y = true_value, shape = cens)) +
  geom_point(alpha = 0.3) +
  scale_y_continuous(limits = c(-0.9, 1)) +
  geom_hline(yintercept = log(LOD), linetype = 'dashed') +
  scale_shape_manual(values = c(16,2)) +
  xlim(-4, 5) +
  theme_minimal() +
  theme(plot.caption=element_text(hjust = 0), legend.position = c(0.77, 0.15), legend.background = eleme
    fill = NA, size =0.5, linetype = "solid"),
        panel.border = element_rect(colour = "black", fill=NA, size=0.5),legend.title = element_blank()
  labs(x='X', y = 'Log(Y)', caption= "Figure 1b: The same data of 1000
observations is plotted without censoring any of
the observations.") +
  geom_text(aes(x = 4.5, y = 0.05), label = "LOQ", size = 3, show.legend = FALSE, family = "serif") +
  labs(x = 'X', y = 'Y')

grid.arrange(censored_plot, uncensored_plot, ncol = 2)
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```

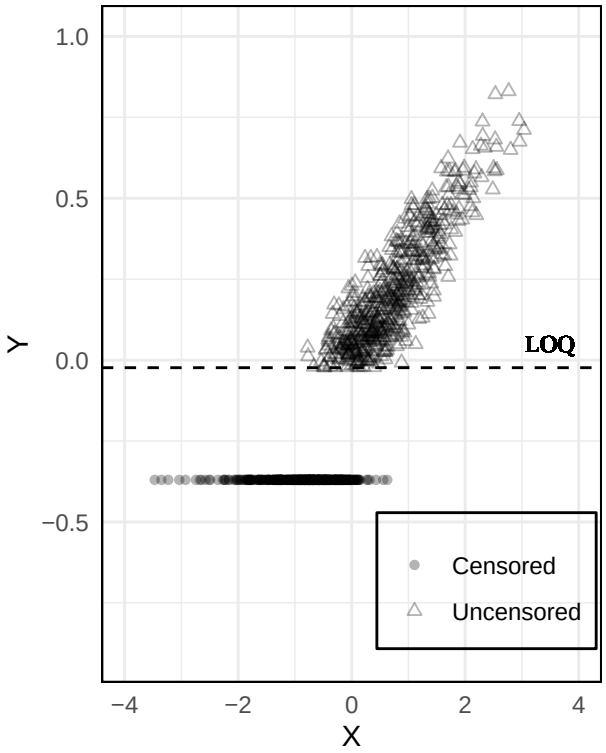

Figure 1a: Example of 1000 observations following a log−normal distribution. Censored observations are replaced with the LOQ divided by the squareroot of 2

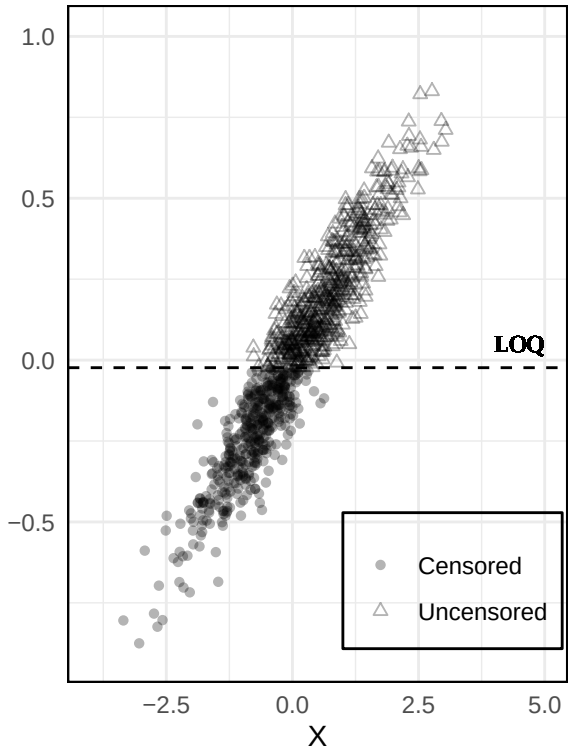

Figure 1b: The same data of 1000 observations is plotted without censoring any of the observations.

```
ggsave(plot = grid.arrange(censored_plot, uncensored_plot, ncol = 2), dpi = 300, file = "Example_Censor
```

```
## Saving 6.5 x 4.5 in image
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```