Homework_1

Zekai Jin

2022-09-19

Loading packages

```
library(tidyverse)
```

Problem 1

a) qualitative_ordinal
b) qualitative_binary
c) qualitative_nominal
d) quantitative_continuous
e) quantitative_discrete

Problem 2

a)

```
bike_data = c(45, 39, 25, 47, 49, 5, 70, 99, 74, 37, 99, 35, 8, 59)
summary(bike_data)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    5.00   35.50   46.00   49.36   67.25   99.00
```

```
range(bike_data)[2]-range(bike_data)[1]
```

```
## [1] 94
```

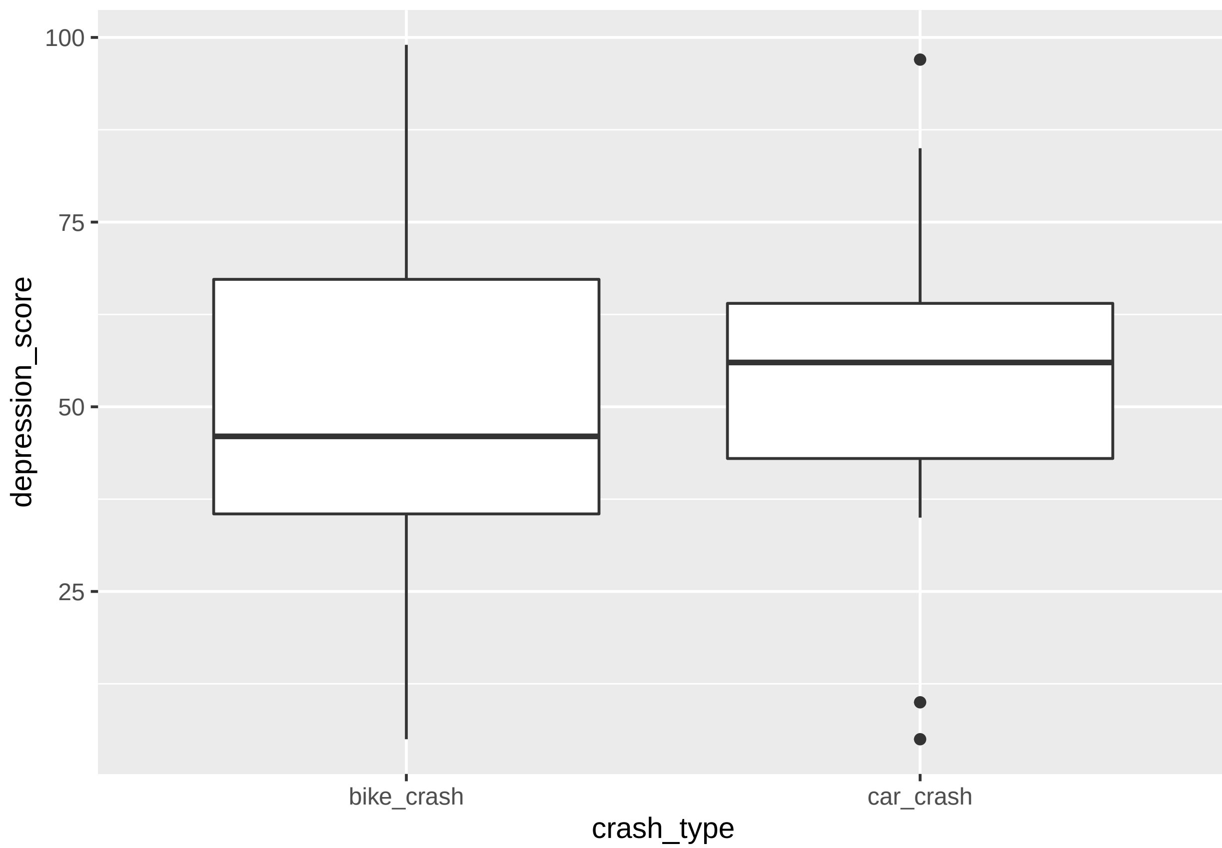
```
sd(bike_data)
```

```
## [1] 28.84603
```

Based on the code above, it is easy to find the mean is 49.36, median is 46, range is 94 and the SD is 28.85. We then make a plot of the depression score in both studies using the following code,

```
car_data=c(67, 50, 85, 43, 64, 35, 47, 97, 58, 58, 10, 56, 5)
combined_df=tibble(depression_score=c(bike_data,car_data),
                   crash_type=factor(c(rep("bike_crash",14),rep("car_crash",13)))
                   )

ggplot(combined_df,aes(crash_type,depression_score))+geom_boxplot()
```



a) b) Thus, the depression score distribution in the bike crash study is mostly symmetric but slightly left-skewed, and the depression score in the car crash is slightly right-skewed. However, from the boxplot, we cannot immediately tell if the distribution is unimodal or not.

c) Comparing the two box plots, individuals with a recent car crash history is more likely to have a higher depression score.

Problem 3

a) $P(A) = \frac{\text{Even Outcomes}}{\text{All Outcomes}} = \frac{6}{12} = \frac{1}{2}$

b) $P(B) = \frac{1}{\text{All Outcomes}} = \frac{1}{12}$

c) Because $A \cap B = \emptyset$, $P(B \cup A) = P(A) + P(B) = \frac{7}{12}$

d) Because $P(B \cup A) = 0 \neq P(A) \times P(B)$, A and B are not independent.

Problem 4

Let A represent "a woman 75 years older have dementia", and B represent "a woman 75 years older have positive CT scan". Obviously, we have

$$P(A) = 0.05$$

$$P(B|A) = 0.8$$

$$P(B|A^c) = 0.1$$

According to Bayes' theorem,

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.8 \times 0.05}{0.8 \times 0.05 + 0.1 \times 0.95} \\ &= 0.2963 \end{aligned}$$