week3assignment_JC

Jessalyn Chuang

2025-01-29

#(1) Discrete Probability Distributions

a. Say you flip a fair, unweighted coin 20 times. What is the probability of obtaining 5 or fewer heads?

```
prob <- pbinom(q = 5, size = 20, prob = 0.50)
print(paste("The probability of obtaining 5 or fewer heads is", round(prob, 3)))
```

```
## [1] "The probability of obtaining 5 or fewer heads is 0.021"
```

b. Now, you switch to an unfair, weighted coin whose probability of landing on heads is 80%. If you flip the new, weighted coin 15 times, what is the probability of obtaining 10 or more heads?

For this question, I calculated the probability of 10 or more heads.

```
# Find P(X >= 10) using the complement rule: P(X >= 10) = 1 - P(X <= 9)
unfair_prob <- 1 - pbinom(q = 9, size = 15, prob = 0.80)
print(paste("The probability of obtaining 10 or more heads is", round(unfair_prob, 3)))
```

```
## [1] "The probability of obtaining 10 or more heads is 0.939"
```

c. Many busy ocean shipping routes overlap with key whale migratory routes along the coast. In the Santa Barbara Channel, the mean number of whale strikes (by shipping vessels) is 12 per year. What is the probability that only 6 whale strikes will occur this coming year? What is the probability that 6 or fewer whale strikes will occur this coming year?

```
prob_only_6 <- dpois(x = 6, lambda = 12)
prob_6_or_fewer <- ppois(q = 6, lambda = 12)

print(paste("The probability of only 6 whale strikes is", round(prob_only_6, 3)))
```

```
## [1] "The probability of only 6 whale strikes is 0.025"
```

```
print(paste("The probability of 6 or fewer whale strikes is", round(prob_6_or_fewer, 3)))
```

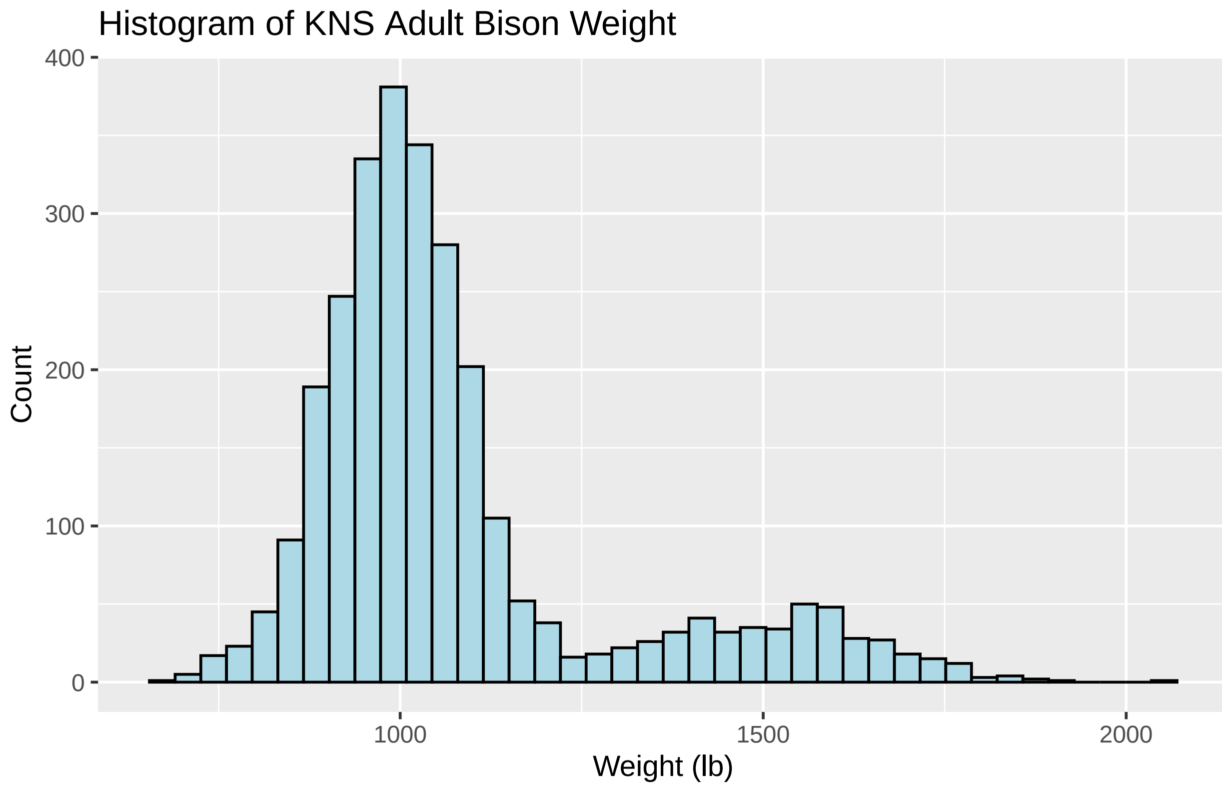
```
## [1] "The probability of 6 or fewer whale strikes is 0.046"
```

#(2) Assessing Normality

a. Each year since 1994, the Konza Prairie Biological Station in Kansas rounds up a herd of 300+ bison to weigh and ID each individual. Bison display sexual dimorphism, meaning males are typically much larger than females. In addition, there is a strong, positive relationship between age (in years) and weight (in pounds). We are interested in seeing whether the weights of adult bison (age > 3 years) are normally distributed.

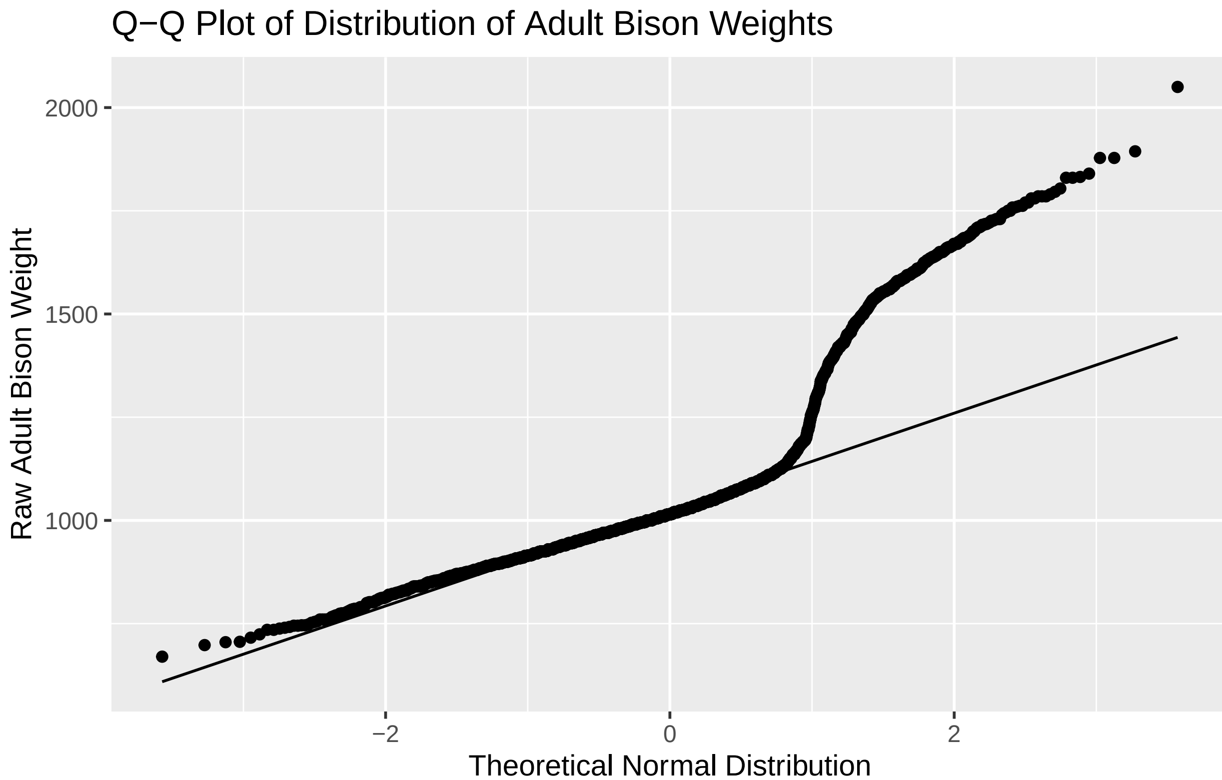
Load in the knz_bison.csv dataset and examine the data using the View() function. Create a new dataset to filter for records only from adult bison (age > 3 years). Then, create a histogram and a quantile-quantile plot examining the distribution of weights of adult bison. Make sure you use an appropriate number of bins in your histogram and that both figures include proper axes labels and captions.

```
knz_bison_data <- read.csv(file = here("./Assignments/week3/knz_bison.csv"),
                           stringsAsFactors = TRUE)

#View(knz_bison_data)

knz_bison_data$age <- knz_bison_data$rec_year - knz_bison_data$animal_yob
knz_bison_filtered <- subset(knz_bison_data, age > 3)

#Number of bins chosen based on Rice rule (# bins = 2n^(1/3))
ggplot(knz_bison_filtered, aes(x = animal_weight)) +
  geom_histogram(bins = 40, fill = "lightblue", color = "black")+
  labs(x = "Weight (lb)",
       y = "Count",
       title = "Histogram of KNS Adult Bison Weight",
       caption = "This histogram shows the distribution of KNS Adult Bison weight.
       The data appears bimodal because it shows two distinct peaks.")
```



This histogram shows the distribution of KNS Adult Bison weight.
The data appears bimodal because it shows two distinct peaks.

```
ggplot(knz_bison_filtered, aes(sample = animal_weight))+
  geom_qq()+
  geom_qq_line()+
  labs(x = "Theoretical Normal Distribution",
       y = "Raw Adult Bison Weight",
       title = "Q-Q Plot of Distribution of Adult Bison Weights",
       caption = "Q-Q plot of adult bison weights, showing strong deviations
       from normality, particularly in the upper tail.")
```



Q−Q plot of adult bison weights, showing strong deviations from normality, particularly in the upper tail.

b. For the same bison data, calculate and report the skewness and kurtosis values for the weight of adult bison.

```
#Skewness
print(paste("Bison Data Skewness:",round(skewness(knz_bison_filtered$animal_weight),2)))
```

```
## [1] "Bison Data Skewness: 1.53"
```

This data is highly skewed as |s| is greater than 1.

```
#Kurtosis
print(paste("Bison Data Kurtosis:",round(kurtosis(knz_bison_filtered$animal_weight),2)))
```

```
## [1] "Bison Data Kurtosis: 4.76"
```

This data has a leptokurtic distribution since k > 3.

c. Based on the information provided, your figures, and your analyses, explain in 1-2 sentences why you would or would not assume these data are normally distributed.

I would not assume the data are normally distributed. The histogram suggests bimodality, the Q-Q plot shows significant deviations from normality, and the distribution appears skewed with leptokurtic characteristics.

#(3) Continuous Probability Distributions

a. Continuing to use the same bison data, create a new dataset that filters your adult bison data only for female bison. What are the mean and standard deviation values for weights of adult, female bison?

```
knz_bison_female = subset(knz_bison_filtered, animal_sex == "F")
mean_female = mean(knz_bison_female$animal_weight)
print(paste("The mean weight for female bison is", round(mean_female,2), "lbs."))
```

```
## [1] "The mean weight for female bison is 992.58 lbs."
```

```
sd_female = sd(knz_bison_female$animal_weight)
print(paste("The standard deviation of weight for female bison is", round(sd_female,2), "lbs."))
```

```
## [1] "The standard deviation of weight for female bison is 90.01 lbs."
```

b. Assuming the adult, female bison weight data are normally distributed, what is the probability of capturing an adult, female bison weighing less than 900 pounds? Weighing between 900 and 1100 pounds? What is the value below which 95% of adult female bison are expected to weigh?

```
#female bison weighing less than 900 pounds
female_bison_less_900 <- pnorm(q = 900, mean = mean_female, sd = sd_female)
print(paste("The probability of capturing an adult, female bison weighing less than 900 pounds is", rou
```

```
## [1] "The probability of capturing an adult, female bison weighing less than 900 pounds is 0.15"
```

```
#female bison weighing between 900 and 1100 pounds
female_bison_less_1100 <- pnorm(q = 1100, mean = mean_female, sd = sd_female)
female_bison_900_1100 <- female_bison_less_1100 - female_bison_less_900
print(paste("The probability of capturing an adult, female bison weighing 900-1100 pounds is", round(fe
```

```
## [1] "The probability of capturing an adult, female bison weighing 900-1100 pounds is 0.73"
```

```
#value below which 95% of adult female bison are expected to weigh
print(paste("The weight below which 95% of adult female bison are expected to weigh is", round(qnorm(p =
```

```
## [1] "The weight below which 95% of adult female bison are expected to weigh is 1140.63 lbs."
```