HW 3 - Beimnet Taye

2023-02-13

P1

1

-Distributions allow us to better visualize the probability of certain events that random variables assign. CDFs, PDFs, and quantile functions all allow us to directly quantify probabilities of events without us having to get the pre-image of the RV's output to derive the same information.

P2

1

```r
# TALBE
  values <- tibble(w = seq(0,1,length = 1000),
                   X = w^2,
                   Y = (1-w)^2,)

# CDF FUNCTION
  CDF <- function(df, rv) {

    rv <- enquo(rv)
    df %>%
      arrange(!!rv) %>%
      mutate(p = row_number()/nrow(df)) %>%
      ggplot() +
      geom_line(aes(x = !!rv, y = p))

  }
# X CDF
  X_CDF <- CDF(values,X)
  X_CDF
```
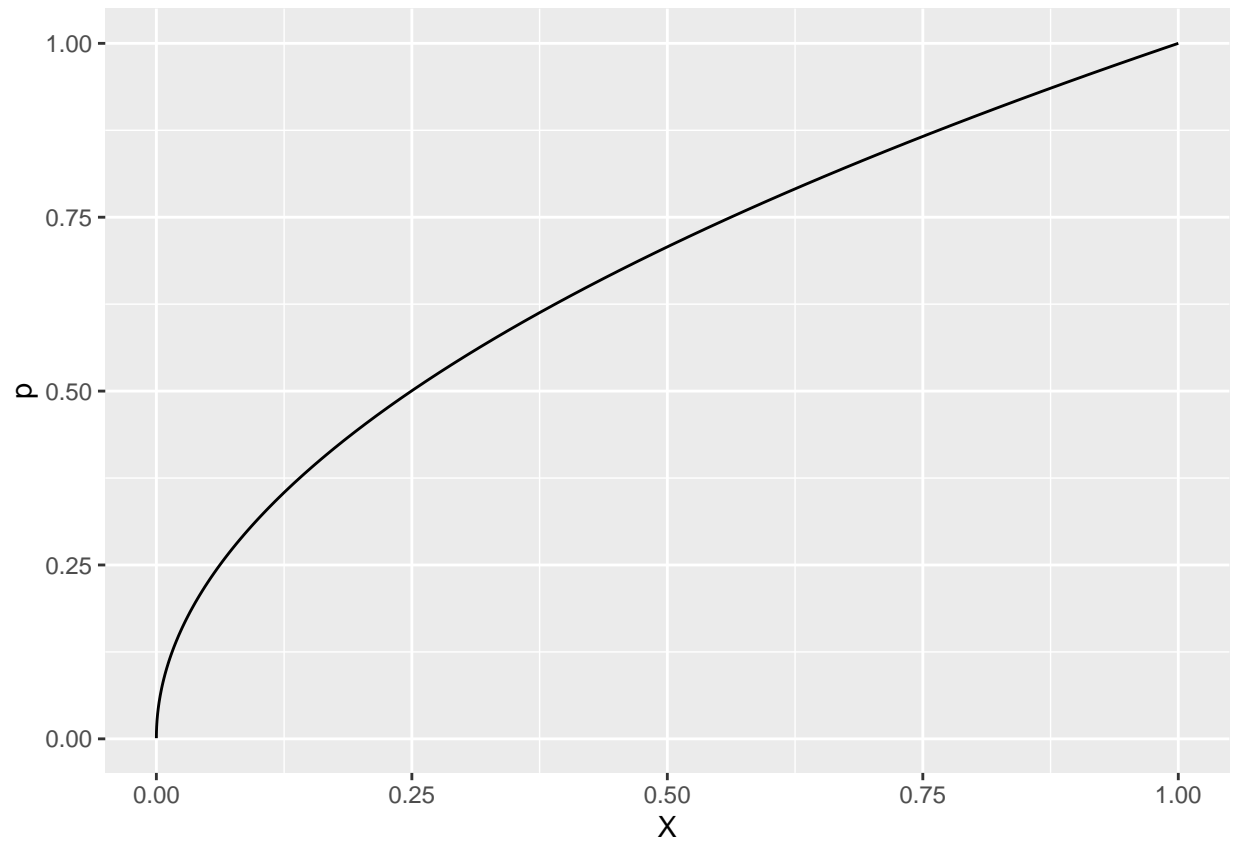
```
# X PDF
  X_PDF <- values %>%
    ggplot()+
    geom_histogram(aes(x = X), binwidth = .01)

  X_PDF
```
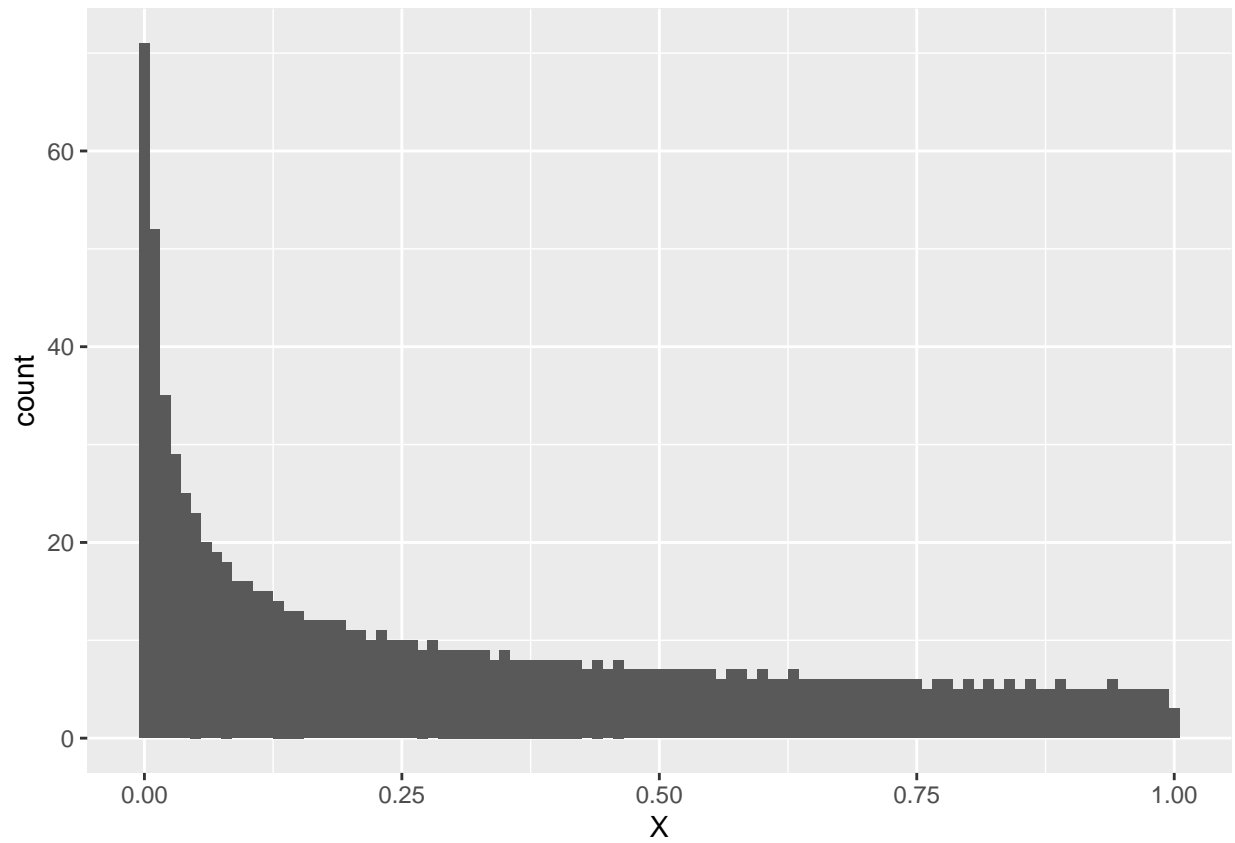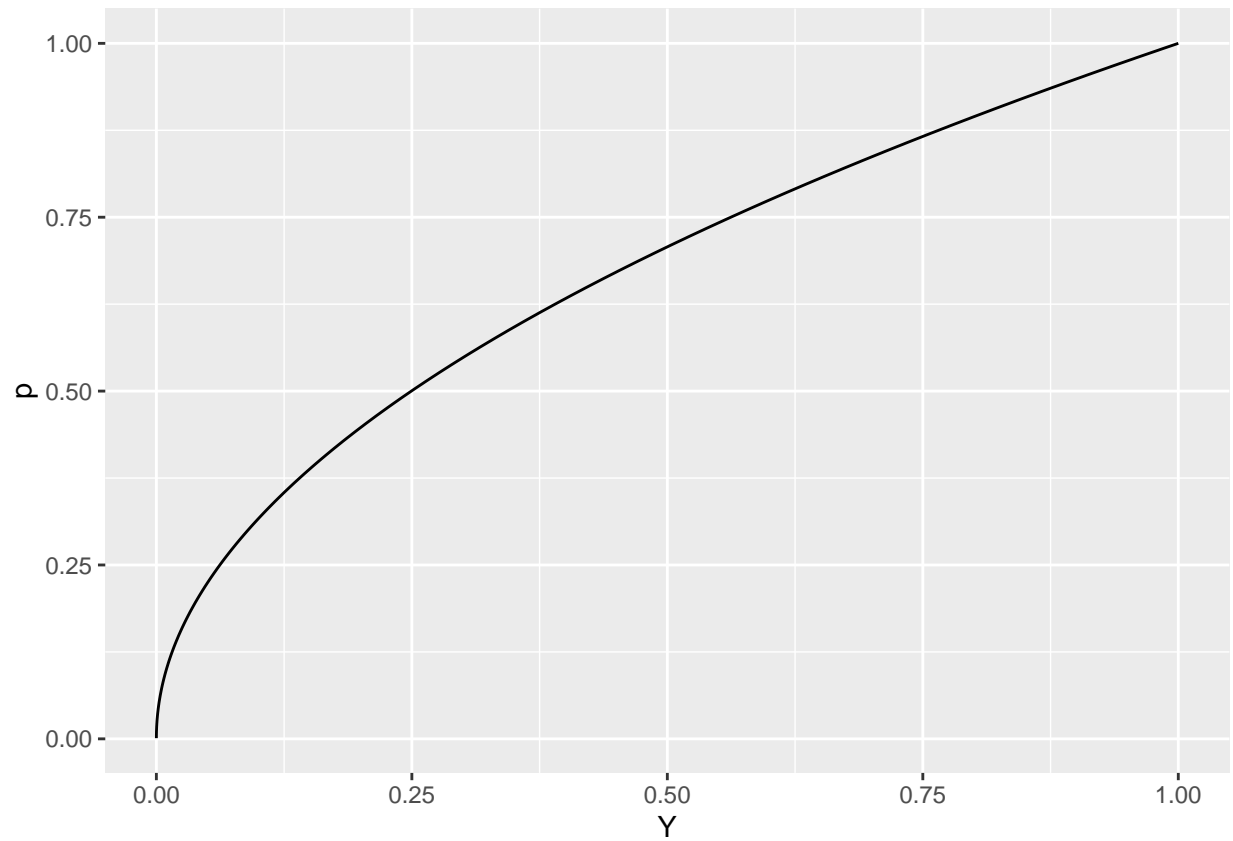
```
  # Y CDF
Y_CDF <- CDF(df=values,rv = Y)

Y_CDF
```
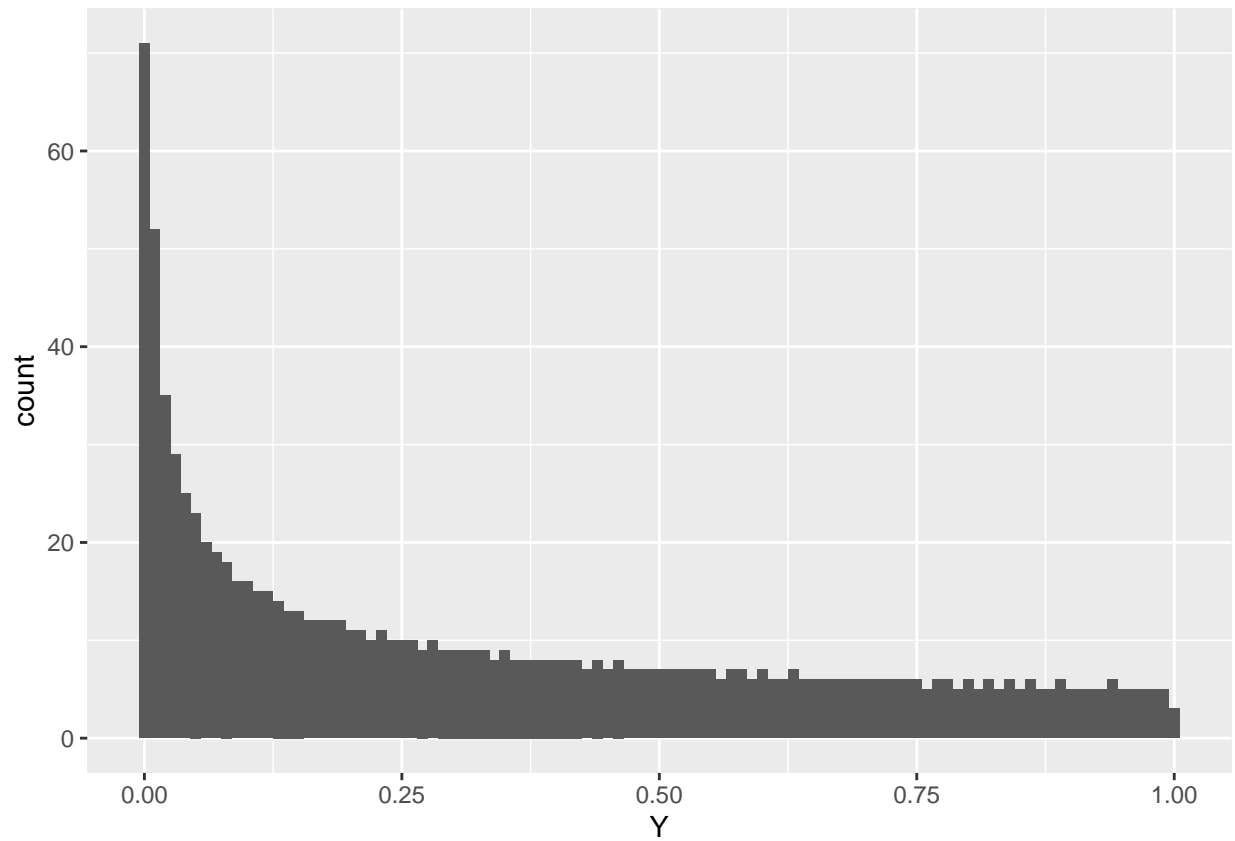
```
# Y PDF

Y_PDF <- values %>%
  ggplot()+
  geom_histogram(aes(x = Y), binwidth = .01)

Y_PDF
```
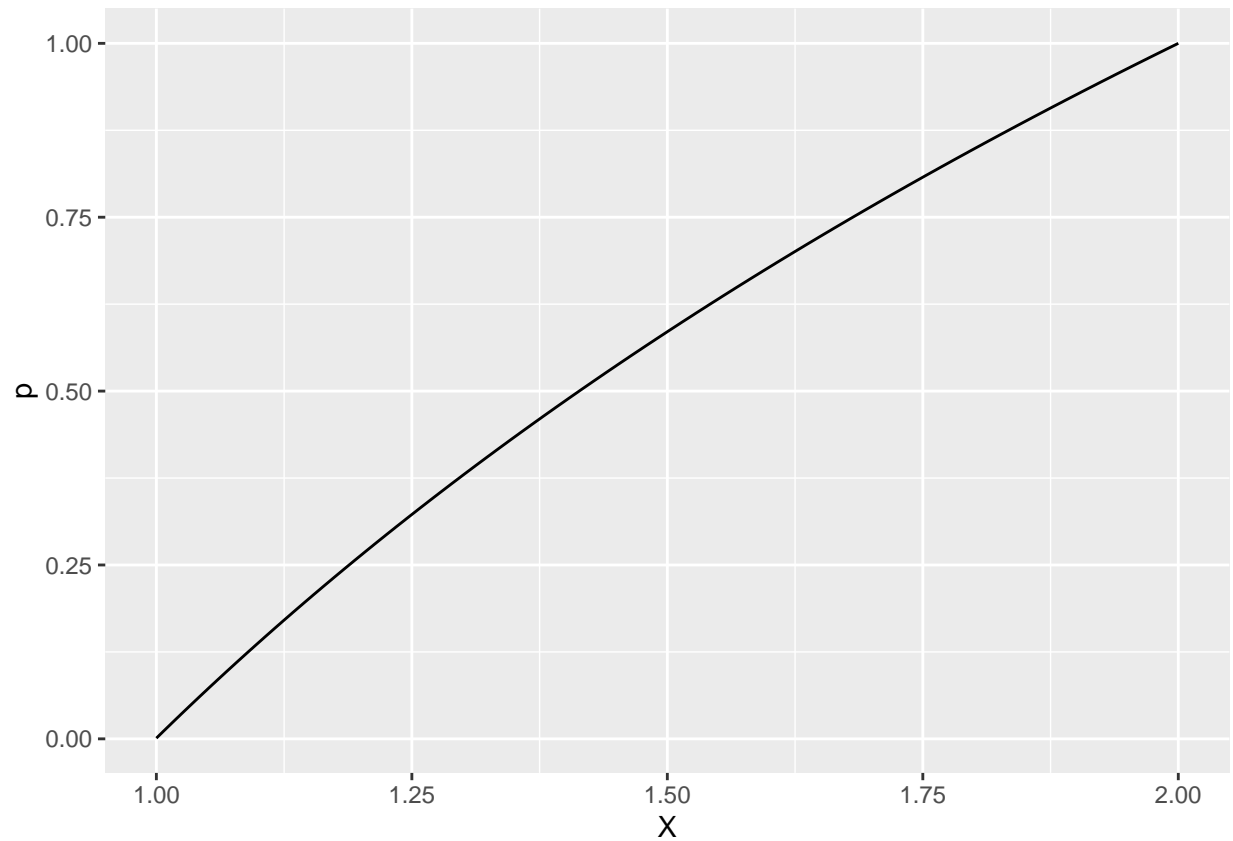
2 & 3

```
#TABLE
custom <- tibble(w = seq(0,1,length = 1000),
                 X = 2^w,
                 Y = 2^(1-w))
# CUSTOM X CDF
CX_CDF <- CDF(custom,X)

CX_CDF
```
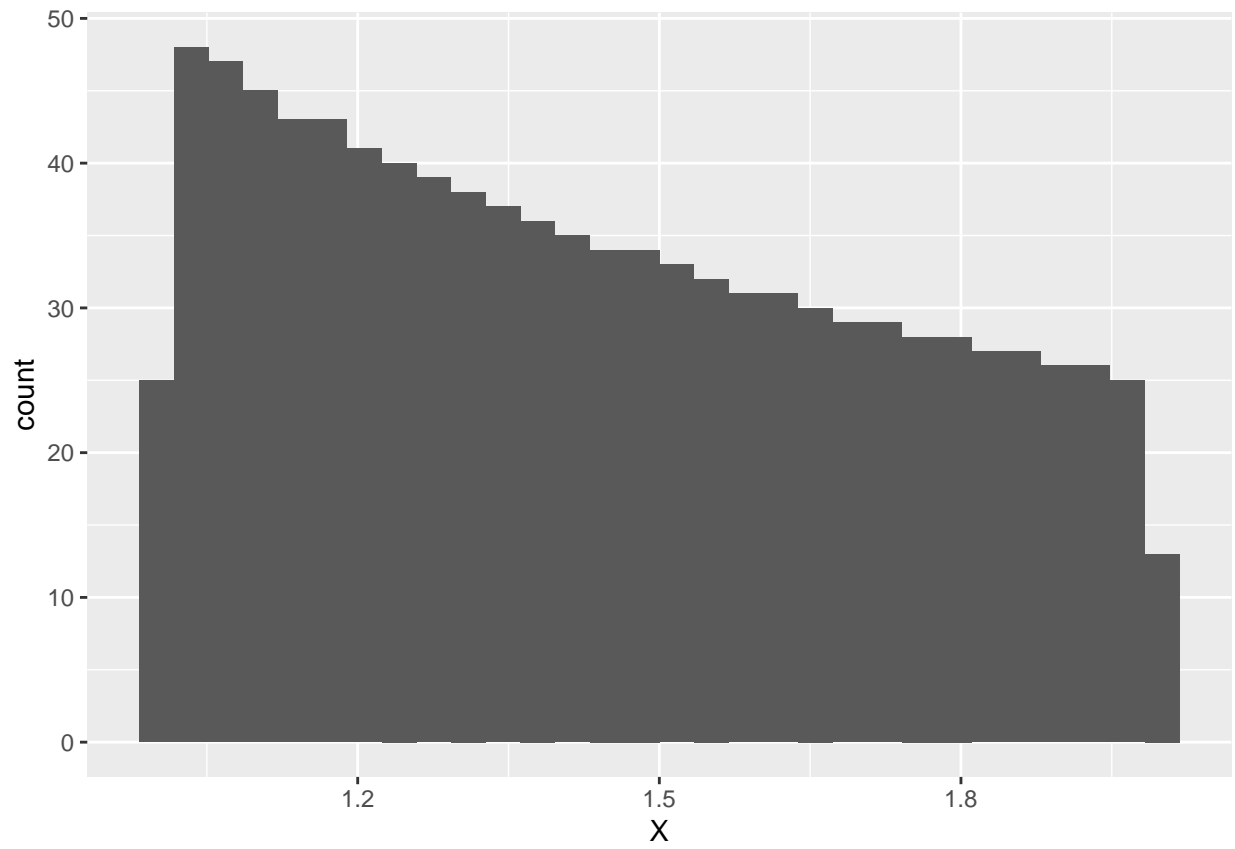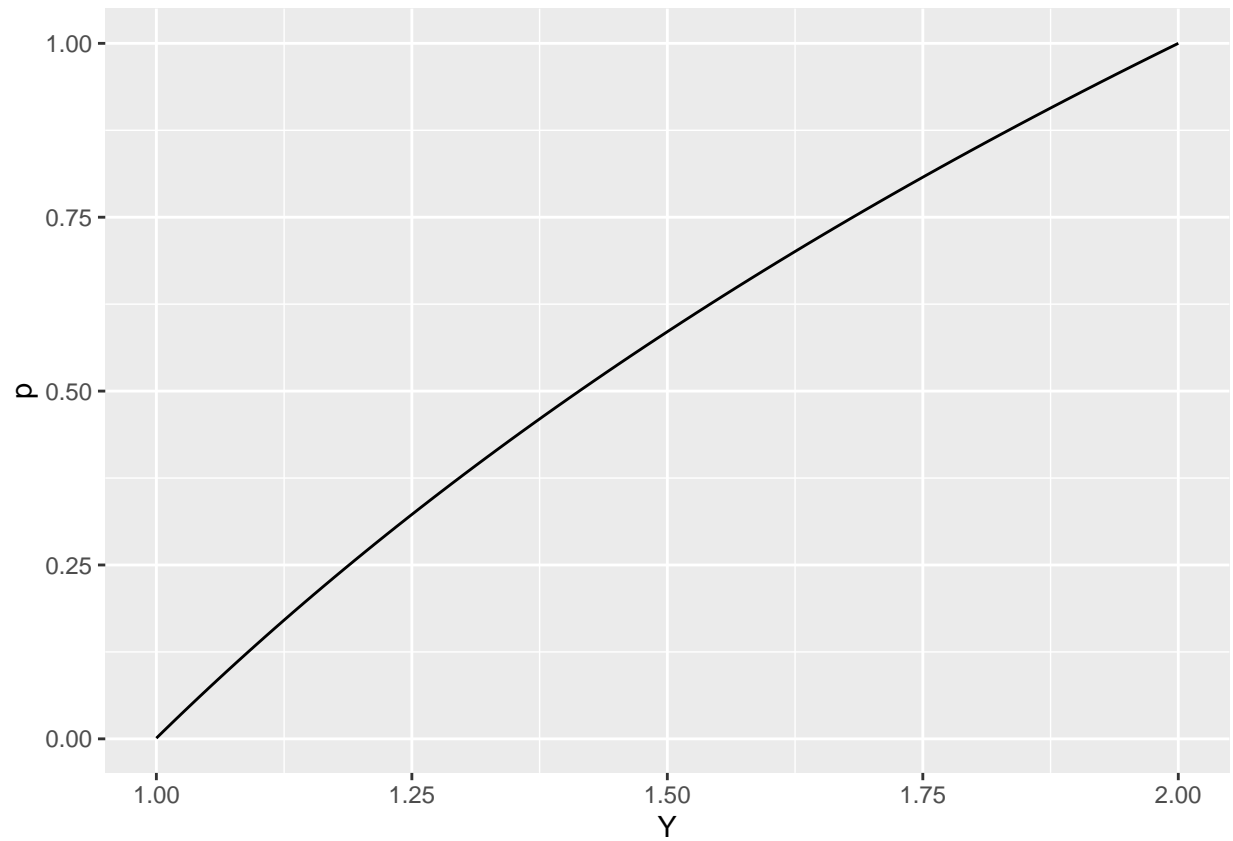
```
# CUSTOM X PDF

CX_PDF <- custom %>%
  ggplot()+
  geom_histogram(aes(x = X))

CX_PDF
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
# CUSTOM Y CDF
CY_CDF <- CDF(custom,Y)

CY_CDF
```
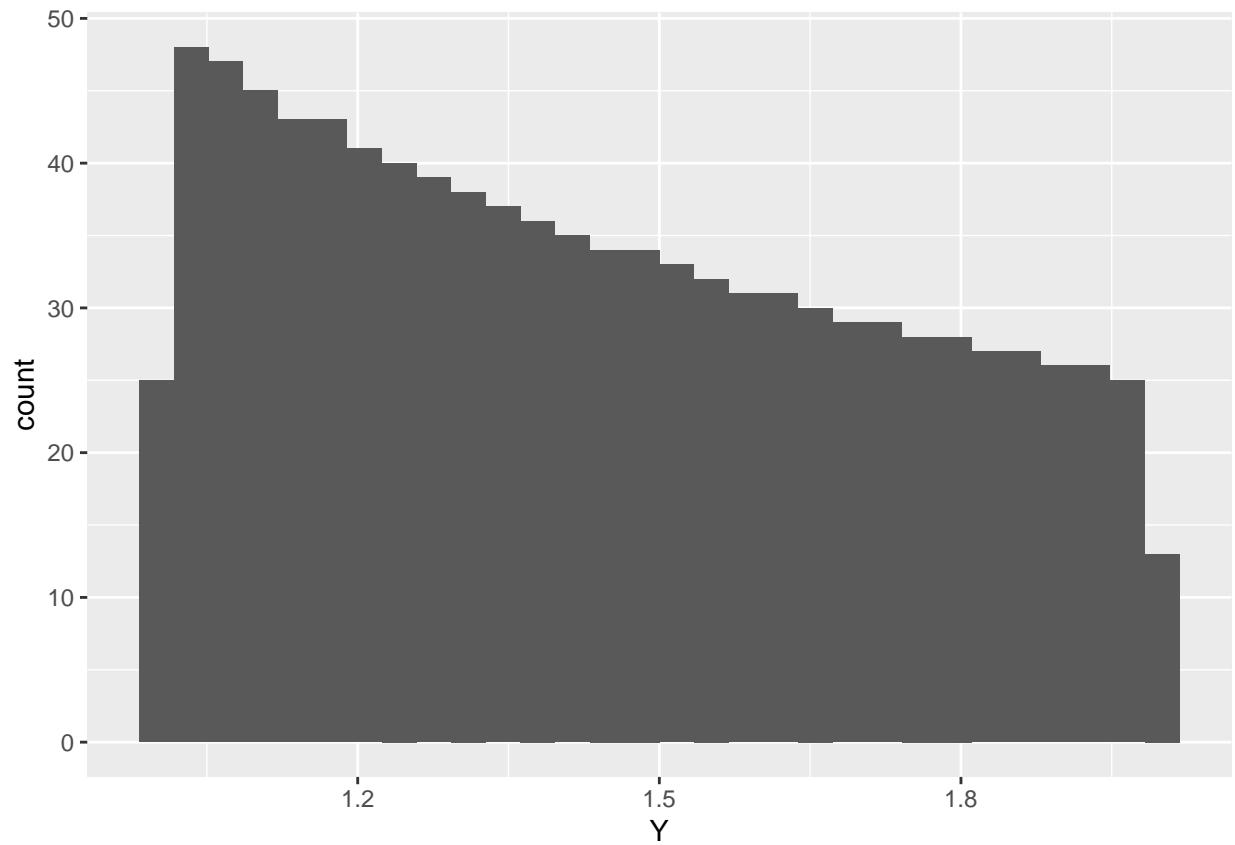
```
# CUSTOM Y PDF
CY_PDF <- custom %>%
  ggplot()+
  geom_histogram(aes(x = Y))

CY_PDF
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

P3

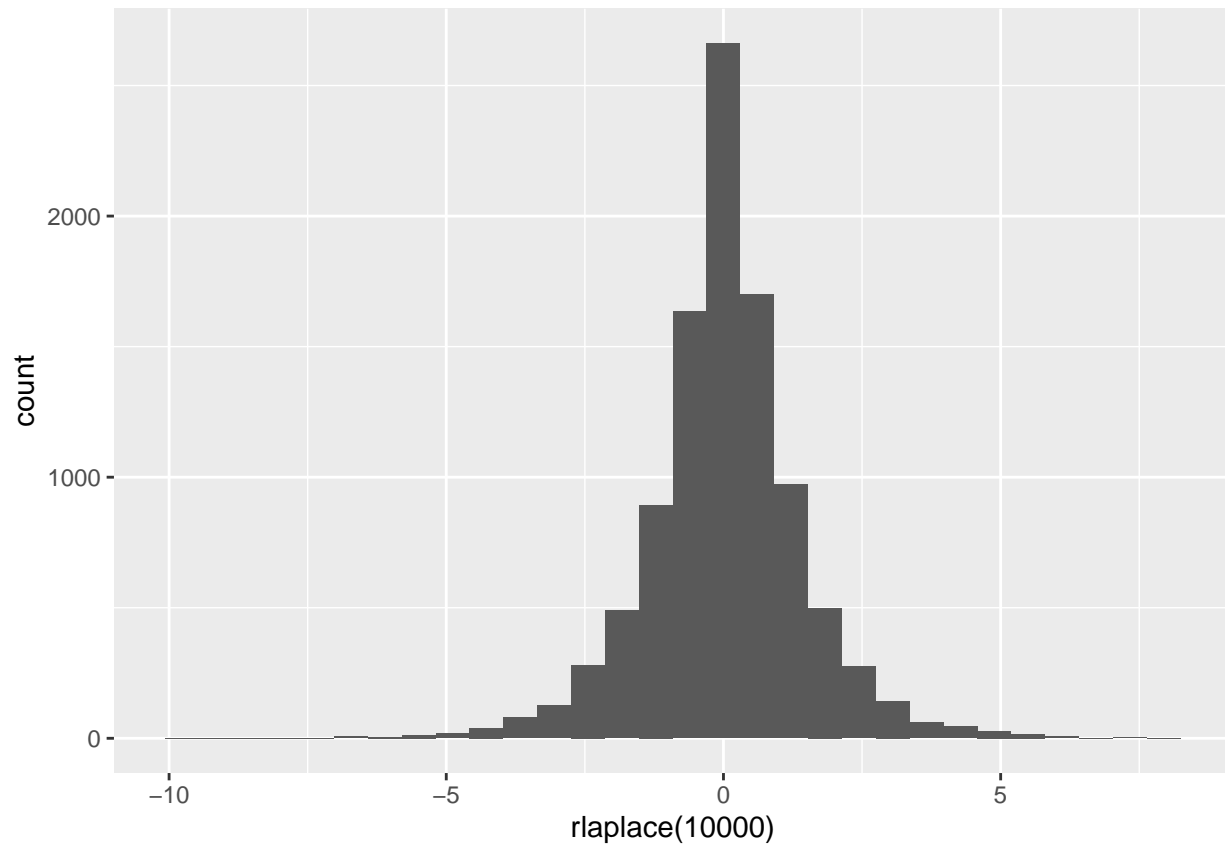**1**

```
# quantile function
qlap <- function(p, a = 0, b = 1) {
  return(case_when(0 <= p & p < 0.5 ~ a+b*log(2*p),
          p >= 0.5 & p <= 1 ~ a-b*log(2*(1-p)),
      )
  )
}
# Sampling function
rlaplace <- function(n,location = 0,scale = 1) {
  tibble(p = runif(n),
        X = qlap(p,location,scale)) %>%
    pull(X)
}
```
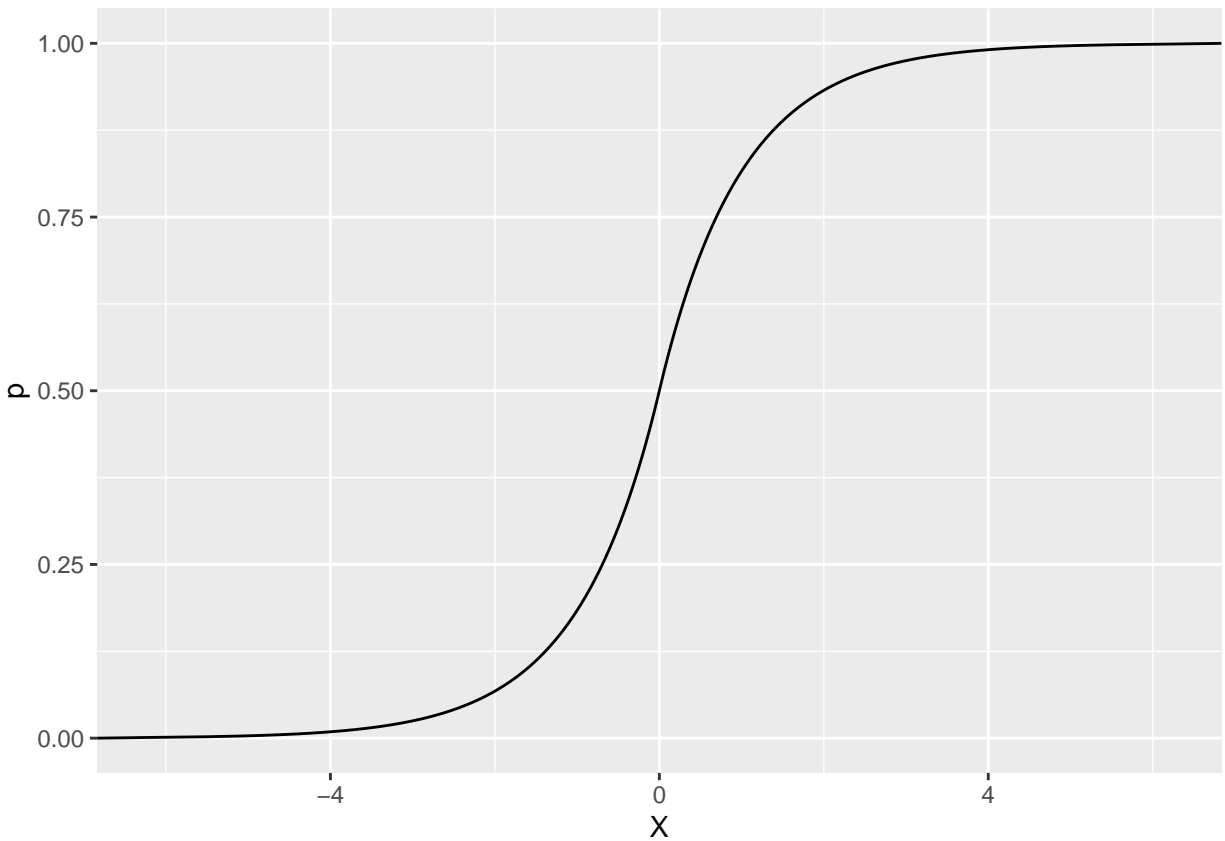
2

```
lap_PDF <- ggplot() +
  geom_histogram(aes(x = rlaplace(10000)))
lap_PDF
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.


```
lap_CDF <- tibble(p = seq(0,1,length = 1000),
             X = qlap(p)) %>%
  ggplot() +
  geom_line(aes(x = X, y = p))
lap_CDF
```

P4

1

$-P(Y > -102) = \frac{1}{3}$

2

$-P(0.5 < X < 1) = 0.95 - 0.25 = 0.7$

P5

1

-Height in a given population is probably normally distributed as individual height will symmetrically cluster around the population mean and symmetrically taper off in frequency on both sides as the height measures get more extreme.

2

-A fair coin toss has a Bernoulli distribution as it is a discrete distribution with two possible outcomes: heads or tails, with a probability of 0.5 for either outcome.

3

-The number of work place accidents at a construction site in a year probably has a poisson distribution. The number of workplace accidents is a countable event within a given time interval, is relatively rare, and the probability of the occurrence of one accident is independent of the probability of the occurrence of another.

4

-The probability of rolling a particular value on a six sided die is uniformly distributed as each value has an equal probability of occurrence (1/6).

5

-The amount of time it takes to receive an order of chicken nuggets from a particular McDonalds location is probably exponentially distributed as the wait times will cluster around relatively small values (it's just chicken nuggets) and dramatically fall off from there in an exponential fashion.