Rather than write a project for sampling distributions, we get this instead. A review for test 1 that can be done with the computer.

a distribution described both a data set, and a population. The same words applied to both, although the definitions are different. For example, the sample mean of a data set is $(x_1 + x_2 + \cdots x_n)/n$, whereas the mean of a discrete random variable is $\sum_k k(P(X = k)$.

For each, find the mean.

Question 1: Find the mean of the following with `mean()`

```
2,35,8,13,17,19
```

Question 2: Run the following command and estimate the mean from the histogram

```
> hist(rchisq(100, 20))
```

Question 3: A discrete random variable has distribution

```
k       |  0    1    2   3
=================================
P(X=k) | 1/2   1/3  1/8 ?
```

Find the mean.

Question 4: The following graphs the density of a distribution. Find the mean from the graph.

```
> curve(dchisq(x, 20), 0, 50)
```

The standard deviation measures spread.

Question 5: Find the standard deviation of the data set

```
1 3 2 4 1 45
```

Question 6: Find the standard deviation of the random variable X with $P(X = 1) = p$, $P(X = 0) = 1 - p$.

Question 7: Estimate the standard deviation from this randomly chosen graph

```
> curve(dnorm(x, sd = runif(1)), -2, 2)
```

The median also measures center.

Question 8: Find the median of the data set

```
  1 3 2 7 7 6 3 2 10 10,000,000
```

Question 9: Create the following graphic and estimate the median and mean

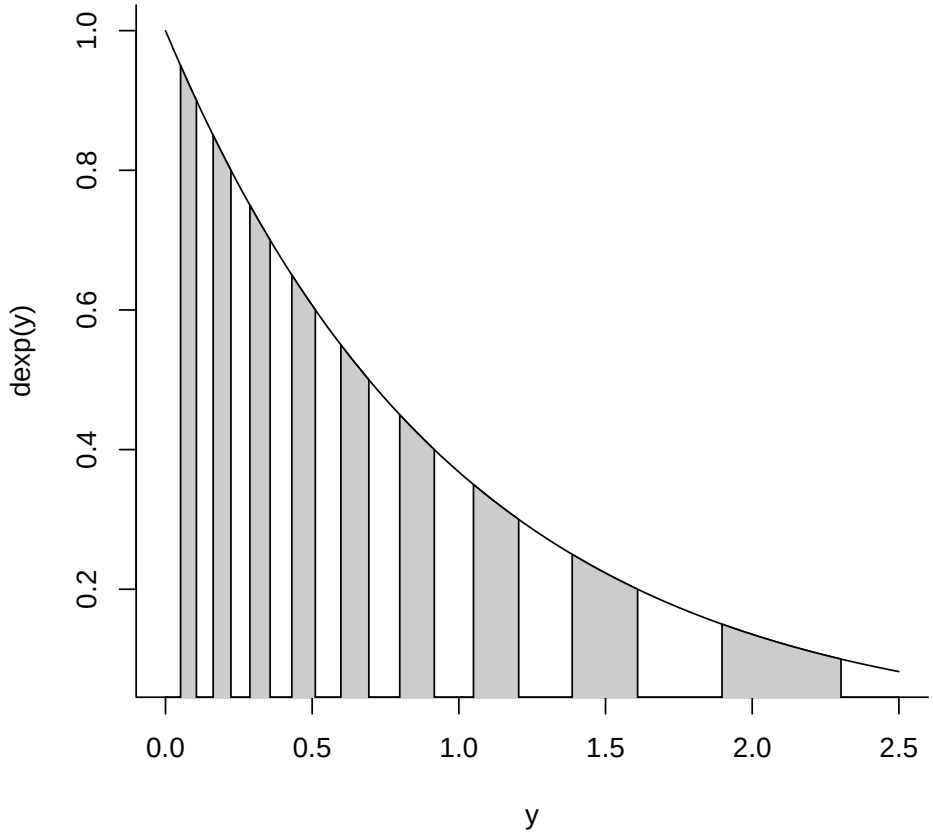
```
> hist(rf(100, 10, 10))
```

The percentiles or quantiles split the data by percentages. Q_1 and Q_3 are the 25th and 75th percentile.

Question 10: Find Q_1 and Q_3 for the data set

```
3 5 6 2 4 10 12
```

Question 11: Explain the shaded figure.



Question 12: Find min, max, IQR and estimate the median and mean from the graph made by

```
> boxplot(rlnorm(100, 2, 1))
```

Question 13: Make the following graphs, estimate the correlation, then check with `cor()`

```
> x = rnorm(100)
> y = rnorm(100)
> y1 = rnorm(100, x)
```

```
> y2 = rnorm(100, x, 5)
> plot(x, y)
> plot(x, y1)
> plot(x, y2)
```

Question 14: Load the data set `mtcars` and find the regression coefficients for `mpg` modeled by `wt`.

Load the data with

```
> data(mtcars)
> attach(mtcars)
```

unload with

```
> detach(mtcars)
```

The normal and binomial distributions were our basic named distributions

Question 15: If X is binomial with $n = 10$, $p = .3$ what is the expected value of X, $P(X > 5)$, $P(X = 2)$?

Question 16: Suppose the number of snowy days per year is historically binomial with $n = 100$ and $p = .05$. What is the mean number of snowy days? How likely is there to be no snowy days? How likely is there to be more than 10 snowy days?

Question 17: Let Z be a standard normal, X normal with mean 10 and variance 20. Which is more $P(Z < -2)$ or $P(X < -2)$? What value of x has $P(X \leq x) = .75$?

Question 18: A population of fish is each year is estimated to be normally distributed with mean 10,000 and variance 2,000. What is the probability there are fewer than 9000 fish in a given year?